Application and Bootstrapping of the Munich Chain Ladder Method

February 13, 2016

Author:
Victor SUNDBERG
Supervisor:
Boualem DJEHICHE

Abstract

Point estimates of the Standard Chain Ladder method (CLM) and of the more complex Munich Chain Ladder method (MCL) are compared to real data on 38 different datasets in order to evaluate if MCL produces better predictions on average with a dataset from an arbitrary insurance portfolio. MCL is also examined to determine if the future paid and incurred claims converge as time progresses. A bootstrap model based on MCL (BMCL) is examined in order to evaluate its possibility to estimate the probability density function (PDF) of future claims and observable claim development results (OCDR). The results show that the paid and incurred predictions by MCL converge. The results also show that when considering all datasets MCL produce on average better estimations than CLM with paid data but no improvement can be seen with incurred data. Further the results show that by considering a subset of datasets which fulfil certain criteria, or by only considering accident years after 1999 the percentage of datasets in which MCL produce superior estimations increases. When examining BMCL one finds that it can produce estimated PDFs of ultimate reserves and OCDRs, however the mean of estimate of ultimate reserves does not converge to the MCL estimates nor do the mean of the OCDRs converge to zero. In order to get the right convergence the estimated OCDR PDFs are centered and the mean of the BMCL estimated ultimate reserve is set to the MCL estimate by multiplication.

Keywords: Claim reserving, Munich Chain Ladder (MCL), Bootstrap

Sammanfattning

Punktskattningar gjorda med Standard Chain Ladder (CLM) och den mer komplexa Munich Chain Ladder-metoden (MCL) jämförs med verklig data för 38 olika dataset för att evaluera om MCL ger bättre prediktioner i genomsnitt än CLM för en godtycklig försäkringsportfölj. MCLs prediktioner undersöks också för att se om de betalda och de kända skadekostnaderna konvergerar. En bootstrapmodell baserad på MCL (BMCL) undersöks för att utvärdera om möjligheterna att estimera täthetsfunktionen (probability density function, PDF) av framtida skadekostnader och av "observable claim development results (OCDR)". Resultaten visar att MCLs estimerade betalda och kända skadekostnader konvergerar. Resultaten visar även att när man evaluerar alla dataseten så ger MCL i genomsnitt bättre prediktioner än CLM med betald data, men ingen förbättring kan ses med CLM med känd skadekostnadsdata. Vidare visar resultaten även att genom att bara titta på dataset som uppfyller vissa krav, eller genom att bara använda olycksår efter 1999, så ökar andelen dataset där MCL ger bättre prediktioner än CLM.Vid evaluering av BMCL ser man att den kan producera estimerade PDF:er för ultimo-reserver och OCDR:er, men att medelvärdet av ultimo-reserv prediktionerna från BMCL inte konvergerar mot MCL-prediktonerna och att medelvärdet av OCDR:erna inte konvergerar mot noll. För att få rätt konvergens så centreras OCDR PDF:erna och ultimo-reservernas medelvärden sätts till mostsvarande MCL-prediktionens värde genom multiplikation.

Svensk Titel: Om Bootstrapping Av Munich Chain Ladder

Nyckelord: Reservsättning, Munich Chain Ladder (MCL), Bootstrap

Acknowledgment

The author is most grateful to Marina Ann Stolin at Trygg-Hansa as she is the person who suggested the subject of the thesis and who during the writing of the thesis provided important feedback and contacts within Trygg-Hansa. The author would like to thank Trygg-Hansa and Codan for allowing their data to be used in the analysis. The feedback and expert opinions of Malcolm Cleugh, Rasmus Hemström, Randi Langstrup Pedersen and Tine Hvithamar was greatly appreciated.
The author would like to thank supervisor professor Boualem Djehiche at KTH for quick and insightful feedback. A thanks also to the authors girlfriend for help with proofreading and adding the final touches. Finally the author would like to thank his family for their never ending support and encouragement.

Contents

Chapter 1

Introduction

The main idea behind insurance is that people can pool their risks and thereby drastically decrease the risk for each individual, meaning for example that each individual does not need to have enough capital to buy a new house if their house burns down. By paying a premium the individual gets covered by the insurance and if something happens the individual receives compensation.

The insurance company does not need to have enough capital at hand to cover all claims if every risk they are exposed to occurs at once, as the probability of that happening is near zero and it would defeat the entire point of pooling the risks together. However, the individual taking an insurance policy needs to know that if something happens the insurance company has enough reserve capital to pay his claim. This means that the insurance company needs to have enough reserve capital that the likelihood of it not being able to pay future claims to their insurance policy holders is extremely low, while not requiring so much capital that the advantages of pooling the risks disappear.

When calculating the estimated total claim cost for an insurance portfolio the most commonly used method is the Standard Chain Ladder method (CLM). CLM can use either paid or incurred claims to estimate the total claim cost. A problem is that the estimation of the total claim cost can differ by a large margin depending on if the paid or the incurred data is used. An idea is that by applying a method that uses both paid and incurred data one can improve the estimation of the total claim cost. One method that uses both the paid and incurred data is the Munich Chain Ladder method (MCL) (Mack and Quarg (2004)).

There are several different methods based on CLM that can be used to estimate future probability density functions (PDF) of claims and claim reserves. These methods include the over-dispersed Poisson (ODP) bootstrap model (Renshaw and Verrall (1998) and England (2002)) and Mack's Model (Mack (1993)) with a distribution assumption. MCL is a newer and more complex method which does not have as much research into it as CLM. As such it does not have as many methods to estimate claim and reserve PDFs, however there are some. One is a blockwise bootstrapped method that is proposed in (Liu and Verrall (2010)).

In the papers where MCL and Liu-Verrall method are presented they are applied to datasets with no missing data and datasets which fit the methods well. In this thesis the methods are applied to 38 different datasets from 38 different insurance portfolios with different times to settle claims, different number of accident years and some with missing data, faulty data and/or problematic parametric values. The questions that are asked are; can MCL decrease the difference in future reserve estimations between the predictions made with paid and incurred claims, can MCL give an improvement of accuracy in estimating the total claims cost over CLM and can an estimated PDF of reserves be made with a method based on MCL.

Chapter 2

Chain Ladder Methods

2.1 Standard Chain Ladder

The Standard Chain Ladder method (CLM) is often used in the insurance industry to calculate the needed reserves as it is a simple method which gives an estimation of the future reserves. With CLM one approximates the factor describing how cumulative paid or incurred (paid plus claim reserve) claims grows over a development period, often from one year to the next.

2.1.1 Method

CLM assumes that one can make a good approximation of future cumulative paid or incurred claims by setting:

$$C^P_{i,j+1} = C^P_{i,j} \cdot f^P_j \qquad \text{and} \qquad C^I_{i,j+1} = C^I_{i,j} \cdot f^I_j \tag{2.1}$$

where $C^P_{i,j}$ is the cumulative paid claims and $C^I_{i,j}$ is the cumulative incurred claims for accident period i after j periods of development, f^P_j and f^I_j are two scalars, the paid and incurred development factor for development period j.

From Eq. (2.1):

$$C^P_{i,J} = C^P_{i,j} \prod_{k=j}^{J-1} f^P_k \qquad \text{and} \qquad C^I_{i,J} = C^I_{i,j} \prod_{k=j}^{J-1} f^I_k \tag{2.2}$$

given that $C_{i,j}$ and $f_j \dots f_J$ are known and $J > j$

Claim Triangle and Development Factor

Before approximating the development factor a claim triangle is needed. To set up a claim triangle one starts by looking at an accident period, $i = 1$, J periods ago, which has had J periods to develop. The next accident period, $i = 2$, has had $J-1$ development periods, and this goes on until accident period J which has only had one period to develop. By putting all the cumulative claims for each accident period and development period into a $J \times J$ matrix where the rows are accident periods and columns are development periods one creates a claim triangle, see Table 2.1 for an example.

The development factor is given by:

$$f_j^{\mathbf{K}} = \frac{\sum_{n=1}^{J-j} C_{n,j+1}^{\mathbf{K}}}{\sum_{n=1}^{J-j} C_{n,j}^{\mathbf{K}}} \tag{2.3}$$

where $J - j$ is the number of accident years which has had $j+1$ periods of development. $\mathbf{K}= \{P, I\}$

	1	2	3	4	$J=5$
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$		
4	$C_{4,1}$	$C_{4,2}$			
$J=5$	$C_{5,1}$				

Table 2.1: $J = 5$. Cumulative claim triangle with the accident years as rows and the development years as columns. $C_{i,j}$ is the cumulative paid or incurred claim for accident year i after j development years. The calendar year is $i + j - 1$.

Predicting Future Claims

With Eq. (2.1) and (2.3) a claim triangle can be filled and thereby predicting the cumulative claim for each accident periods development period up to development period J, see Table 2.2.

	1	2	3	4	$J=5$
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	$C_{2,4} \cdot f_4$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,3} \cdot f_3$	$C_{3,3} \cdot f_3 \cdot f_4$
4	$C_{4,1}$	$C_{4,2}$	$C_{4,2} \cdot f_2$	$C_{4,2} \cdot f_2 \cdot f_3$	$C_{4,2} \cdot f_2 \cdot f_3 \cdot f_4$
$J=5$	$C_{5,1}$	$C_{5,1} \cdot f_1$	$C_{5,1} \cdot f_1 \cdot f_2$	$C_{5,1} \cdot f_1 \cdot f_2 \cdot f_3$	$C_{5,1} \cdot f_1 \cdot f_2 \cdot f_3 \cdot f_4$

Table 2.2: Same as Table 2.1 but with filled in predictions of claims for future development of the accident periods

2.2 Advanced Chain Ladder Methods

2.2.1 Assumptions of CLM

Despite being a central part of loss reserving in the insurance industry CLM has several weaknesses because of the strong assumptions of the model. These assumptions are:

1. The development of the claim is not dependent on the year of the accident.
2. The average development factor is a good estimator of the future claim development.
3. To use known paid or incurred claims gives a good estimate of the development of the claim.

These strong assumptions are made to make the CLM as simple and straight forward as possible at the possible cost of accuracy in its predictions. There are several more advanced Chain Ladder methods (CL) that use weaker assumptions and include more variables in order to improve the accuracy of the predictions at the cost of simplicity.

2.2.2 Examples of Advanced Chain Ladder Methods

The separation method (Verbeek (1972) and Taylor (1977)) is an advanced CL, which includes the concept of claim inflation. With claim inflation the separation method uses both calandar years and development year when estimating future claims, this means that assumption 1. is thereby replaced by a weaker assumption:

- The development of the claim is dependent on the year of the accident in a way that can be modelled.

Another advanced CL is the double chain ladder method (Miranda et al. (2012) and (2015)) which considers two triangles, the paid claims triangle, Table 2.1, and a triangle consisting of the number of reported claims.

A third advanced CL is the Paid Incurred Chain Claims Reserving Method (Merz and Wüthrich (2010)) or the PIC method. This method works by considering both the paid and incurred data and forcing the cumulative paid and incurred values for the final development year to be identical. This method weakens assumption 3:

- Using known paid and incurred claims together simultaneously gives a good estimate of the development of the claim.

2.3 The P/I-Problem with Standard Chain Ladder

In this section the P/I-divergence, a problem given by assumption 3. is examined. From now on $P_{i,j}$ and $I_{i,j}$ will be used interchangeably with $C^P_{i,j}$ and $C^I_{i,j}$ respectively.

2.3.1 P/I-ratio

The P/I-ratio, $Q(i,j)$, is the ratio of the paid and incurred claim:

$$Q(i,j) = \frac{C^P_{i,j}}{C^I_{i,j}} \tag{2.4}$$

As the incurred claims are the paid claims plus the claim reserve, R^{claim}, the equation can be rewritten as:

$$Q(i,j) = \frac{C^P_{i,j}}{C^P_{i,j} + R^{claim}_{i,j}} \tag{2.5}$$

An assumption that can be made is that R^{claim} should rarely, if ever, be negative as it is meant to be a reserve and not a loan. Another assumption that can be made is that as development time progresses the cumulative paid claims grows and the needed reserve becomes smaller. Those assumptions together with Eq. (2.5) give:

- $Q(i,j) \leq 1.$
- $\lim_{j\to\infty} Q(i,j) = 1$

In the following subsection the P/I-ratio will be examined in both real data and in predictions from CLM.

P/I-ratio In Real Data and CLM

The datasets, dataset A and B, that are presented in this section are from (Mack and Quarg (2004)) and (Merz and Wüthrich (2010)) respectively. The development periods in all datasets in this thesis are development years.

The real data in both graphs in Fig. 2.1 follows the conclusions from 2.3.1 as neither have a $Q(i,j)$ greater than 1 and in both datasets $Q(i,j)$ goes toward 1 as the development time increases. However looking at the predictions from CLM it is clear that their Q-values neither converge nor stay below 1.

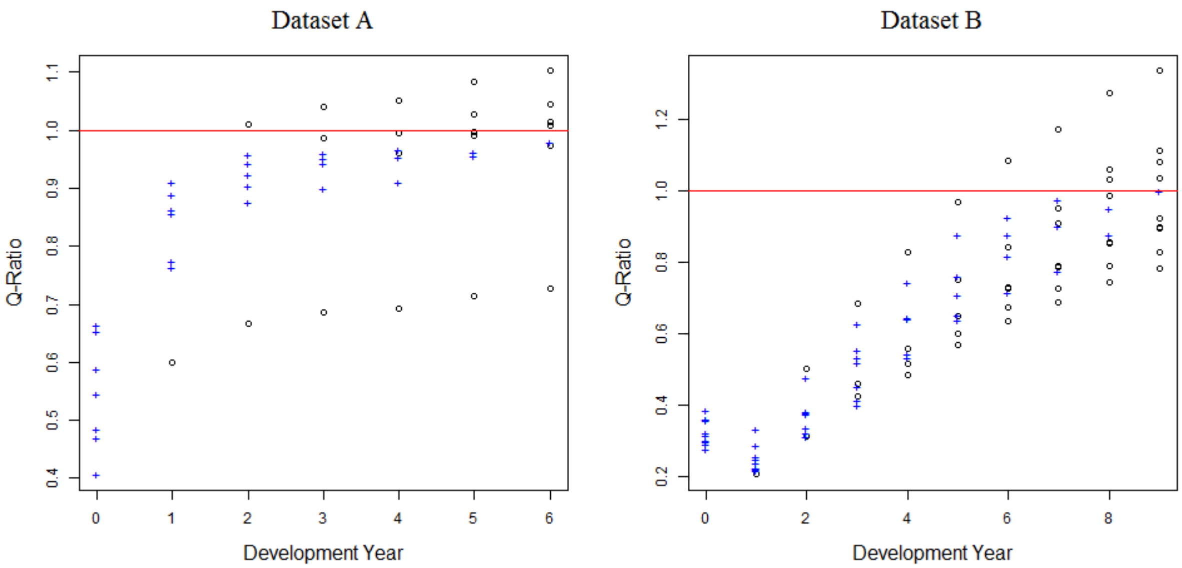

Figure 2.1: Datasets with real life paid and incurred claims. Development year on the X-axis and the Q-ratio on the Y-axis. Blue "+" are real claim data, black "o" are predicted claim data. The red line is at the Q-ratio equal to 1 and as can be seen, all real data points are below it.

2.3.2 Correlation Between Paid and Incurred Claims

One possible explanation for some of the difference in the Q-ratio pattern could be that CLM does not take into consideration any correlation between the claim triangles that might exist. In this subsection two correlations will be examined: one between Q-ratios and incurred individual development factors and one between I/P-ratios and paid individual development factors. The individual development factors are defined as:

$$F_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}$$

In order to get enough data points to be able to see if there is an underlying correlation and not just the random nature of real life data, this thesis will consider all development years at the same time instead of considering one development year at the time. In order to evaluate all the development years at the same time one can consider the standardized residuals instead of the values, see section 2.4.1.

Fig. 2.2 clearly shows that there is a significant correlation in dataset A, especially between the individual paid development factor and the I/P-ratio, however Fig. 2.3 on the other hand shows that dataset B does not have a correlation nearly as strong.

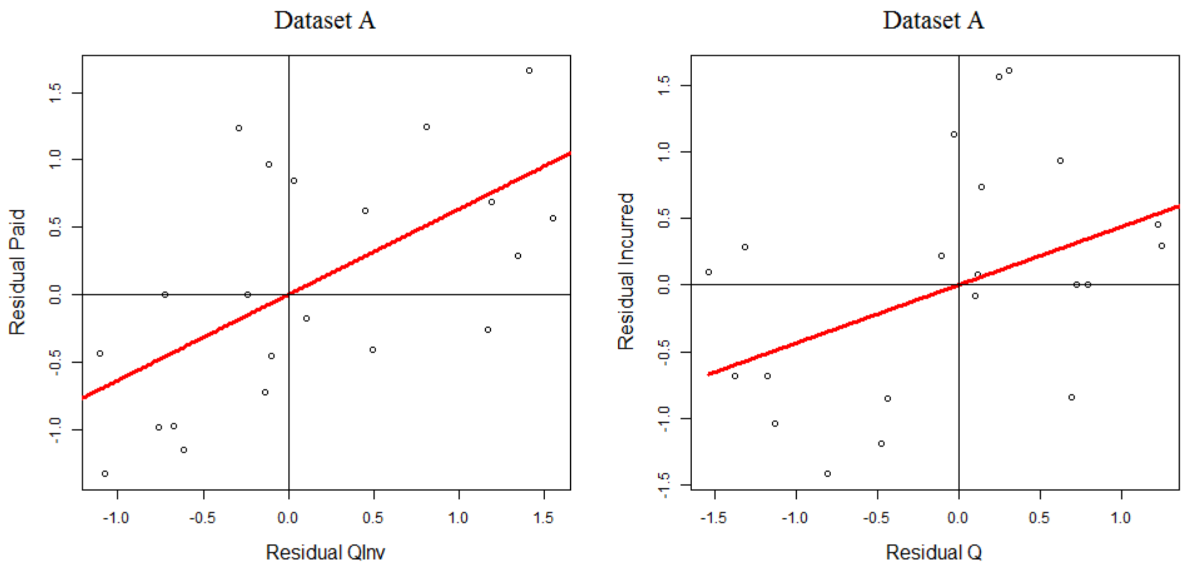

Figure 2.2: Dataset A. Both the X- and Y-axis are the standardized residuals. The correlation between Q^{-1} and F^P is 0.6151 and the correlation between Q and F^I is 0.4415. The two red lines are linear regressions without a constant with $\lambda^P = 0.6360$ and $\lambda^I = 0.4362$ respectively

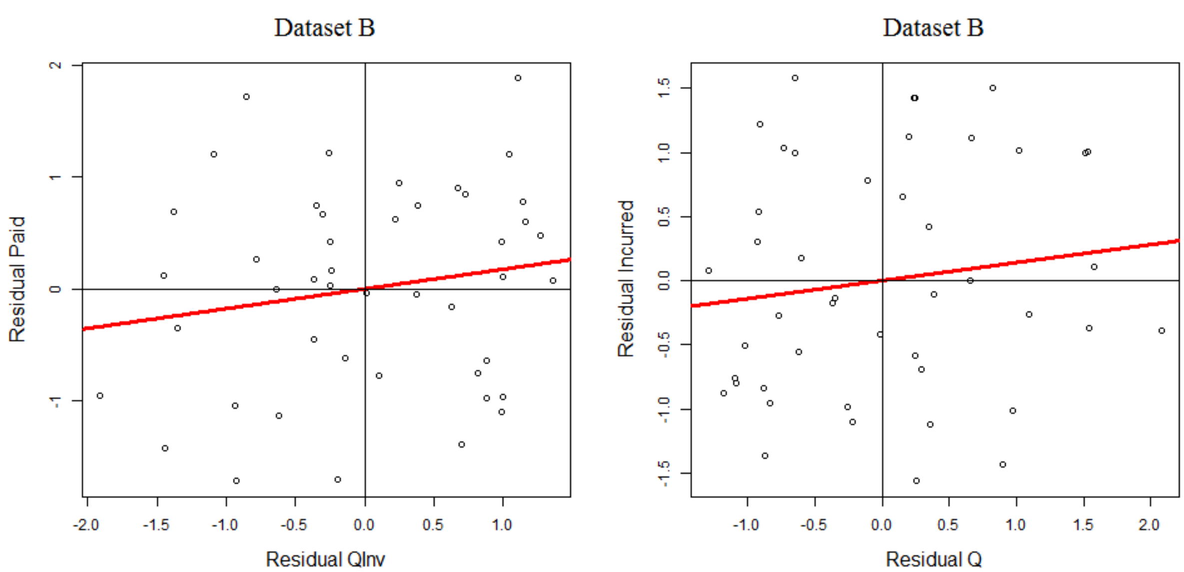

Figure 2.3: Dataset B. Both the X- and Y-axis are the standardized residuals. The correlation between Q^{-1} and F^P is 0.1704 and the correlation between Q and F^I is 0.1365. The two red lines are linear regressions without a constant with $\lambda^P = 0.1764$ and $\lambda^I = 0.1404$ respectively

2.4 The Munich Chain Ladder Method

The Munich chain ladder method (MCL) is an advanced Chain Ladder method that uses both paid and incurred data simultaneously in order to better follow the Q-pattern that the real life data shows. MCL was suggested by Mack and Quarg and all theory and equations in this section is taken from (Mack and Quarg (2004)).

2.4.1 Definitions

Sigma-Algebras

The following three sigma algebras will be used:
$\mathbf{P}_i(s) = \sigma\{P_{i,1}, ..., P_{i,s}\}$, $\mathbf{I}_i(s) = \sigma\{I_{i,1}, ..., I_{i,s}\}$ and $\mathbf{B}_i(s) = \sigma\{P_{i,1}, ..., P_{i,s}, I_{i,1}, ..., I_{i,s}\}$

Residual

The standardized conditional residual of a stochastic variable X and a sigma algebra $\mathbf{B}$ is defined as:

$$\mathbf{Res}(X|\mathbf{B}) = \frac{X - \mathbf{E}(X|\mathbf{B})}{\sqrt{Var(X|\mathbf{B})}}$$

Linear Regression Without a Constant

With two datasets $\bar{X}$ and $\bar{Y}$, $\hat{\lambda}$ is defined as the $\hat{\lambda}$ that minimizes the sum of square errors, $\sum \theta^2$ in equation

$$\bar{X} = \hat{\lambda}\bar{Y} + \theta$$

2.4.2 Method

There are several assumptions used in MCL, the first one is the same as Eq. (2.3) used in CLM, rewritten with expected values:

$$\mathbf{E}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_i(j)) = f_j^P \qquad \text{and} \qquad \mathbf{E}(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_i(j)) = f_j^I \tag{2.6}$$

One also assumes that there exists proportionality constants $\sigma_j^P \geq 0$ and $\sigma_j^I \geq 0$ (Mack (1993)) such that:

$$\mathbf{Var}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_i(j)) = \frac{(\sigma_j^P)^2}{P_{i,j}} \qquad \text{and} \qquad \mathbf{Var}(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_i(j)) = \frac{(\sigma_j^I)^2}{I_{i,j}} \tag{2.7}$$

These assumptions gives that a higher cumulative claim has a lower variance in its individual development factor.

The same ideas are used for Q and Q_{Inv}, with $q(j)$ and $q_{Inv}(j)$ as conditional expected values and σ_j^Q and $\sigma_j^{Q_{Inv}}$ as proportionality constants.

A third assumption is an independence assumption:

- $\{P_{1,j}\}...\{P_{h,j}\}$ are independent and $\{I_{1,j}\}...\{I_{h,j}\}$ are independent.

Where h is the number of accident years that have j development years.

MCL Individual Development Factors

So far none of the assumptions have said anything about the correlation between the paid and the incurred claim triangle. The correlation between the triangles are given by the fourth assumption:

$$E[\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{B}_j] = f^P_{MCL}(i,j) = f^P_j + \lambda^P \cdot \frac{\sigma(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)}{\sigma(Q^{-1}_{i,j}|\mathbf{P}_j)} \cdot (Q^{-1}_{i,j} - q_{inv}(j)) \tag{2.8}$$

$$E[\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{B}_j] = f^I_{MCL}(i,j) = f^I_j + \lambda^I \cdot \frac{\sigma(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_j)}{\sigma(Q_{i,j}|\mathbf{I}_j)} \cdot (Q_{i,j} - q(j)) \tag{2.9}$$

where f^P and f^I are from Eq. (2.6), $\sigma(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)$ and $\sigma(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_j)$ are from Eq. (2.7), $\sigma(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)$ and $\sigma(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_j)$ are the standard deviations of F^P_i and F^I_i and finally λ^P and λ^I are the slopes of the regression lines without a constant for F_P, Q^{-1} and F_I, Q respectively.

Eq. (2.8) and (2.9) can be simplified by:

$$\frac{\sigma(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)}{\sigma(Q^{-1}_{i,j}|\mathbf{P}_j)} = \frac{\sqrt{\mathbf{Var}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)}}{\sqrt{\mathbf{Var}(Q^{-1}_{i,j}|\mathbf{B}_j)}} = \frac{\sqrt{\frac{(\sigma^P_j)^2}{P_{i,j}}}}{\sqrt{\frac{(\sigma^{Q_{Inv}}_j)^2}{P_{i,j}}}} = \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j} \tag{2.10}$$

where the second equal sign is given by Eq. (2.7)

By doing this and the same for Eq. (2.9) one gets:

$$E[\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{B}_j] = f^P_{MCL}(i,j) = f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j} \cdot (Q^{-1}_{i,j} - q_{Inv}(j)) \tag{2.11}$$

$$E[\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{B}_j] = f^I_{MCL}(i,j) = f^I_j + \lambda^I \cdot \frac{\sigma^I_j}{\sigma^Q_j} \cdot (Q_{i,j} - q(j)) \tag{2.12}$$

Eq. (2.11) and (2.12) are not explicitly written in (Mack and Quarg (2004)), however they are used in the practical example.

Eq. (2.11) and (2.12) can be rewritten into:

$$f_{MCL}^{P}(i,j) = f_j^P + \Delta f_{i,j}^P \qquad \text{and} \qquad \Delta f_{i,j}^P = \lambda^P \cdot \frac{\sigma_j^P}{\sigma_j^{Q_{Inv}}} \cdot (Q_{i,j}^{-1} - q_{inv}(j)) \tag{2.13}$$

$$f_{MCL}^{I}(i,j) = f_j^I + \Delta f_{i,j}^I \qquad \text{and} \qquad \Delta f_{i,j}^I = \lambda^I \cdot \frac{\sigma_j^I}{\sigma_j^{Q}} \cdot (Q_{i,j} - q(j)) \tag{2.14}$$

where f_j^P and f_j^I are the ordinary CLM development factors and $\Delta f_{i,j}^P$ and $\Delta f_{i,j}^I$ are terms that are added to the development factors in order to take the two correlations of F^P, Q and F^I, Q_{Inv} into account.

With Eq. (2.11) and (2.12) MCL deviates from the idea that using known paid or incurred claims independently each give good estimates of the development of the claims, and instead suggest that using known paid and incurred claims together gives a good estimate of the development of the claims.

With Eq. (2.11) and (2.12) the claim triangles can be filled.

Different signs of λ^P and λ^I

With Eq. (2.13), Eq. (2.14) and depending on the sign of λ^P and λ^I one can get four different cases, the first two of which are analyzed in (Mack and Quarg (2004)):

1. The first case is when both λs are positive. If $\lambda^P > 0, \lambda^I > 0$ then $\Delta f_{i,j}^P$ and $\Delta f_{i,j}^I$ will have the same sign as $(Q_{i,j}^{-1} - q_{inv}(j))$ and $(Q_{i,j} - q(j))$ respectively, because $\frac{\sigma_j^P}{\sigma_j^{Q_{Inv}}} \geq 0$ and $\frac{\sigma_j^P}{\sigma_j^{Q_{Inv}}} \geq 0$. This means that the accident years with lower than average Q-values get a higher Q-value the next development year and vice versa.

2. The second case is when both λs are zero. If $\lambda^P = \lambda^I = 0$ then $\Delta f_{i,j}^P = \Delta f_{i,j}^P = 0$ which gives $f_{MCL}^P(i,j) = f_j^P$ and $f_{MCL}^I(i,j) = f_j^I$, $\forall$ i and the method is identical to CLM.

The first case is the most common case and is the case MCL was designed for.

MCL was not designed for datasets with $\lambda^P < 0$ and/or $\lambda^I < 0$ and can give strange predictions for these datasets. In section 3.4.1 these cases will be considered.

2.5 MCL Bootstrapping

MCL does only give a point estimate for future claims, however often a point estimate is not enough information and an estimated future PDF is needed. A common way to get an estimated future PDF for a stochastic variable that is defined by one or several stochastic variables is to use a bootstrapping method (Efron (1979) and Efron and Tibshirani (1993))

2.5.1 Bootstrapping Algorithm

Bootstrapping is not a single method, but a broad definition of methods that uses random sampling with replacement. All data and parameters that are created by bootstrapping methods will be denoted by a tilde. The bootstrapping algorithm works in four steps:

1. Manipulate the data, $\mathbf{X} = \{X_1, X_2.....X_n\}$, by functions that have well defined inverses, $F_i(X_i) = Z_i$ resp. $F_i^{-1}(Z_i) = X_i$, and which has an output, Z_i, that can reasonably be assumed to be I.I.D, i.e. $\mathbf{Z} = \{Z_1, Z_2.....Z_n\}$ is I.I.D. $\mathbf{Z}$ is often a set of residuals.
2. Resample with replacement the output $\{Z_1, Z_2.....Z_n\}$.
3. Input the resampled $\widetilde{\mathbf{Z}}$ into the inverse, $F_i^{-1}(Z_k) = \widetilde{X_i}$, to calculate pseudo-random data $\{\widetilde{X_1}, \widetilde{X_2}.....\widetilde{X_n}\}$.
4. Use the pseudo-random data $\{\widetilde{X_1}, \widetilde{X_2}.....\widetilde{X_n}\}$ to calculate $\widetilde{\theta}$, where $\widetilde{\theta}$ is an observed statistic of interest e.g. the mean, a quantile or the standard deviation.

Step 2 to 4 is then done N number of times and with each iteration a $\widetilde{\theta}$ is calculated. With N number of $\widetilde{\theta}$ it is possible to get an empirical PDF of $\widetilde{\theta}$ which as N increases should converge to the true PDF of $\widetilde{\theta}$. The PDF of $\widetilde{\theta}$ can be assumed to be a good approximation of the PDF of the true θ. This mean as N increases the empirical PDF of $\widetilde{\theta}$ becomes a better approximation of the PDF of the true θ. To get a good approximation N is often set to at least 1,000.

Blockwise Bootstrapping

When bootstrapping with several different sets of residuals at the same time one needs to decide if to use a blockwise bootstrap method. The blockwise and non-blockwise methods are identical in all steps of the bootstrapping algorithm except in the resampling, step 2.

If one has two sets of residuals, $\mathbf{Z}$ and $\mathbf{W}$, with a non-blockwise method they are resampled independently of each other. By doing the resampling independently any correlations that existed between them does not carry over to the resampled $\widetilde{\mathbf{Z}}$ and $\widetilde{\mathbf{W}}$.

With a blockwise bootstrap method a new combined set of residuals, $\mathbf{U}$, is defiend as:
$\mathbf{U} = \{(Z_1; W_1), (Z_2; W_2), ...(Z_n; W_n)\}$. $\mathbf{U}$ is then resampled, meaning that there is only one resampling and Z_i is still with W_i for all elements that are resampled. With the blockwise bootstrapping the correlations between the residuals are carried over to the resampled $\widetilde{\mathbf{Z}}$ and $\widetilde{\mathbf{W}}$.

When bootstrapping CLM only one set of residuals are resampled as CLM considers either paid or incurred data. This means that the bootstrapped CLM methods do not use blockwise bootstrapping. An example of a CLM bootstrap method is the over-dispersed Poisson (ODP) bootstrap method (Renshaw and Verrall (1998) and England (2002)).

2.5.2 Liu-Verrall MCL Bootstrapping Method

As MCL uses both paid and incurred claims and uses the F^P, Q_{Inv} and F^I, Q correlations, a more advanced bootstrapping method than the ones that can be used for CLM is needed. One method proposed in (Liu and Verrall (2010)) uses a blockwise bootstrapping of the residuals of the individual development factors, F^P and F^I, the Q-ratios and the Q_{Inv}-ratios.

Method

The Liu-Verrall MCL method does not change the paid or incurred claims data but it re-estimates the parameters (the development factors, the proportionality constants, λ^P and λ^I) in each iteration. By doing a blockwise bootstrapping of all four sets of residuals the correlation between F^P , Q_{Inv} and F^I, Q_{Inv} can be maintained after the resampling for the pseudo-random $\widetilde{F^P}$, $\widetilde{F^I}$, $\widetilde{Q}$ and $\widetilde{Q}^{-1}$. By rewriting the definition of the development factors and conditional expected values of Q and Q^{-1}, see Eq. (2.15) to (2.18), one can create the pseudo-random $\widetilde{f}^P$, $\widetilde{f}^I$, $\widetilde{q}$ and $\widetilde{q}_{Inv}$.

$$f_j^P = \mathbf{E}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_i(j)) = \mathbf{E}(\frac{P_{i,j} \cdot F_{i,j}^P}{P_{i,j}}|\mathbf{P}_i(j)) \tag{2.15}$$

$$f_j^I = \mathbf{E}(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_i(j)) = \mathbf{E}(\frac{I_{i,j} \cdot F_{i,j}^I}{I_{i,j}}|\mathbf{I}_i(j)) \tag{2.16}$$

$$q(j) = \mathbf{E}(\frac{P_{i,j}}{I_{i,j}}|\mathbf{B}_i(j)) = \mathbf{E}(\frac{I_{i,j} \cdot Q_{i,j}}{I_{i,j}}|\mathbf{B}_i(j)) \tag{2.17}$$

$$q_{Inv}(j) = \mathbf{E}(\frac{I_{i,j}}{P_{i,j}}|\mathbf{B}_i(j)) = \mathbf{E}(\frac{P_{i,j} \cdot Q_{i,j}^{-1}}{P_{i,j}}|\mathbf{B}_i(j)) \tag{2.18}$$

With $\widetilde{F^P}$, $\widetilde{F^I}$, $\widetilde{Q}$, $\widetilde{Q}^{-1}$, $\widetilde{f}^P$, $\widetilde{f}^I$, $\widetilde{q}$ and $\widetilde{q}_{Inv}$ $\widetilde{\lambda}^P$ and $\widetilde{\lambda}^I$ can be approximated and the MCL development factors can be calculated using Eq. (2.19) and (2.20)

$$\widetilde{f}_{MCL}^P(i,j) = \widetilde{f}_j^P + \widetilde{\lambda}^P \cdot \frac{\widetilde{\sigma}_j^P}{\widetilde{\sigma}_j^{Q_{Inv}}} \cdot (\widetilde{Q}_{i,j}^{-1} - \widetilde{q}_{Inv}(j)) \tag{2.19}$$

$$\widetilde{f}_{MCL}^I(i,j) = \widetilde{f}_j^I + \widetilde{\lambda}^I \cdot \frac{\widetilde{\sigma}_j^I}{\widetilde{\sigma}_j^Q} \cdot (\widetilde{Q}_{i,j} - \widetilde{q}(j)) \tag{2.20}$$

With $\widetilde{f}_{MCL}^P(i,j)$, $\widetilde{f}_{MCL}^I(i,j)$ and normally distributed error terms the future claims can be predicted:

$$\widetilde{C}^{\mathbf{K}}_{i,j+1} = \mathrm{N}(C_{i,j} \cdot \widetilde{f}^{\mathbf{K}}_{MCL}(i,j), (\widetilde{\sigma}^{\mathbf{K}}_j)^2 \cdot C_{i,j}) \tag{2.21}$$

where $\mathrm{N}(\mu,\sigma^2)$ is a normally distributed variable with a mean value of μ and a variance of σ^2.

With Eq. (2.21) the claim triangles can be filled and the reserves can be estimated. By doing this N number of times an empirical estimated PDF can be made of the needed reserves.

2.5.3 Bootstrapped Munich Chain Ladder Method

The Liu-Verrall MCL method is a good method, however the method lacks a way of dealing with high Sigma-ratios, see section 3.4.2, and how to deal with missing data. Also in the paper where the method is described (Liu and Verrall (2010)) the last σ^p and σ^I, which cannot be calculated and needs to be assumed, are not given any values.

For these reasons a modified version of the Liu-Verrall MCL method will be used in this thesis. This modified method, the Bootstrapped Munich Chain Ladder method (BMCL) can deal with missing data, see section 3.3.1, and has a way of alleviating the problem given by high Sigma-ratio. The last σ^p and σ^I are assumed to be 0.1, as done in (Mack and Quarg (2004)). A final difference between the Liu-Verrall MCL method and BMCL is that BMCL does not add error terms when estimating future claims.

2.6 Calculating Reserves

2.6.1 Ultimate Reserves and Ultimate Reserve Risk

In this thesis the claims are assumed to be settled at the last known development year, J, that is:

$$C^{\mathbf{K}}_{i,J} = C^{\mathbf{K}}_{i,J+1} = ... = C^{\mathbf{K}}_{i,\inf} \tag{2.22}$$

and given this assumption the ultimate reserve, R_∞, can be calculated as:

$$R^{\mathbf{K}}_\infty = \sum_{i=2}^{J} C^{\mathbf{K}}_{i,J} - P_{i,J-i+1} \tag{2.23}$$

Eq. (2.22) is a very strong assumption and another more complex long tail assumption could be used to include the fact that some of the claims are not settled at development year J, however using Eq. (2.22) makes comparing the predictions to real data very simple.

With a defined R_∞ a new variable, r_∞, can be defined as:

$$r_\infty = VaR_{0.995}(R_\infty - E[R_\infty]) \tag{2.24}$$

2.6.2 One Year Reserve Risk

At calendar year J the ultimate reserve is calculated with $\mathbf{D}(J)$, where $\mathbf{D}(J)$ is a sigma algebra that represents all paid and incurred claims known at time J. As time progresses money is paid out and new information is learned. At time J one does not know how much will be paid out next year nor how the estimated ultimate reserve will change, however one can use methods to estimate different $\mathbf{D}_n(J+1)$ and their probabilities. Using these scenarios a new ultimate reserve can be calculated for the same accident years, $[R_\infty^{\mathbf{K}}|\mathbf{D}_n(J+1)]$. The relation between the old and the new reserve can be written as:

$$R_\infty^{\mathbf{K}} = [R_\infty^{\mathbf{K}}|\mathbf{D}_n(J+1)] + \sum_{i=2}^{J}(P_{i,J-i+2} - P_{i,J-i+1}) - \sum_{i=2}^{J}\epsilon_i^{\mathbf{K}}(n) \tag{2.25}$$

where $R_\infty^{\mathbf{K}}$ is the estimated ultimate reserve with the information at time J, $[R_\infty^{\mathbf{K}}|\mathbf{D}_n(J+1)]$ is the ultimate reserve for the same accident years but with the information at time $J+1$, $\sum_{i=2}^{J}(P_{i,J-i+2} - P_{i,J-i+1})$ is the amount paid out during calendar year J and ϵ is the term that corrects for the change in estimation of the ultimate reserve given by the new information.

$[R_\infty^{\mathbf{K}}|\mathbf{D}_n(J+1)]$ can be rewritten as:

$$[R_\infty^{\mathbf{K}}|\mathbf{D}_n(J+1)] = \sum_{i=2}^{J}(C_{i,J+1}^{\mathbf{K}}|\mathbf{D}_n(J+1) - P_{i,J-i+2}) = \sum_{i=2}^{J}(C_{i,J}^{\mathbf{K}}|\mathbf{D}_n(J+1) - P_{i,J-i+2}) \tag{2.26}$$

where $C_{i,J+1} = C_{i,J}$ from Eq. (2.22).

Eq. (2.23), (2.25) and (2.26) give:

$$\sum_{i=2}^{J}(C_{i,J}^{\mathbf{K}} - P_{i,J-i+1}) = \sum_{i=2}^{J}(C_{i,J}^{\mathbf{K}}|\mathbf{D}_n(J+1) - P_{i,J-i+2}) + \sum_{i=2}^{J}(P_{i,J-i+2} - P_{i,J-i+1}) - \sum_{i=2}^{J}\epsilon_i^{\mathbf{K}}(n) \leftrightarrow$$

$$\sum_{i=2}^{J}\epsilon_i^{\mathbf{K}}(n) = \sum_{i=2}^{J}(C_{i,J}^{\mathbf{K}}|\mathbf{D}_n(J+1) - C_{i,J}^{\mathbf{K}}) \tag{2.27}$$

This means that $\epsilon_i(n)$ can be seen as the difference between two predictions of the ultimate loss for an accident year. $\epsilon_i(n)$ is called the observed claim development result (OCDR) (Merz and Wüthrich (2008)). The one year reserve risk, r_1, is defined as (Lauzeningks and Ohlsson (2008)):

$$\sum_{i=2}^{J}\epsilon_i(n) = \epsilon(n) \rightarrow r_1 = VaR_{0.995}(\epsilon) \tag{2.28}$$

Chapter 3

Implementation

In the following sections, both the paid and incurred data is assumed to be complete, with all elements ≥ 0, and being able to be written it in the form of Table 3. This is not always the case with real life data and in section 3.3 ways to adjust the methods to imperfect and missing data will be discussed.

	1	2	3	4	$J=5$
1	$P_{1,1}$	$P_{1,2}$	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	$P_{2,1}$	$P_{2,2}$	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	$P_{4,2}$			
$J=5$	$P_{5,1}$				

	1	2	3	4	$J=5$
1	$I_{1,1}$	$I_{1,2}$	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	$I_{2,1}$	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	$I_{4,2}$			
$J=5$	$I_{5,1}$				

Table 3.1: Complete paid and incurred triangles.

3.1 MCL

3.1.1 Estimating The Parameters

All equations in this section are taken from (Mack and Quarg (2004)). The variables needed to be estimated are: the ordinary CLM development factors (f_j^P and f_j^I), the conditional expected value of Q and Q_{Inv} (q and q_{Inv}), the proportionality constants ($\sigma^p, \sigma^p, \sigma^Q$ and $\sigma^{Q_{Inv}}$) and the slopes of the regression lines without a constant for F_P, Q^{-1} and F_I, Q (λ^P and λ^I).

Development Factors and The Expected Value of Q and Q_{Inv}

The development factors for the paid and incurred claims ($\hat{f}^P$ and $\hat{f}^I$) are estimated by Eq. (2.3) and in an equivalent way $\hat{q}$ and $\hat{q}_{Inv}$ are estimated by :

$$\hat{q}(j) = \frac{\sum_{n=1}^{J-j+1} P_{n,j}}{\sum_{n=1}^{J-j+1} I_{n,j}} \qquad \text{and} \qquad \hat{q}_{Inv}(j) = \frac{\sum_{n=1}^{J-j+1} I_{n,j}}{\sum_{n=1}^{J-j+1} P_{n,j}} \tag{3.1}$$

Proportionality Constants

The proportionality constants for the paid and incurred claims ($\hat{\sigma}^p(j)$ and $\hat{\sigma}^I(j)$) for $j = 1 \dots J-2$ are estimated by:

$$\hat{\sigma}^P(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} P_{n,j}(\frac{P_{n,j+1}}{P_{n,j}} - \hat{f}_j^P)^2} \tag{3.2}$$

$$\hat{\sigma}^I(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} I_{n,j}(\frac{I_{n,j+1}}{I_{n,j}} - \hat{f}_j^I)^2} \tag{3.3}$$

For the penultimate development year, $J-1$, there is only one individual development factor and the proportionality constant cannot be approximated by Eq. (3.2) and (3.3), and therefore needs to be assumed. However, then the difference in Q is often small and the last sigma is not of great importance. In (Mack and Quarg (2004)) the last σ^p and σ^I is simply set to 0.1, and the same is done in this thesis.

In a similar way as $\hat{\sigma}^P(j)$ and $\hat{\sigma}^I(j)$ are approximated, the proportionality constants for Q-ratio and Q_{Inv}-ratio ($\hat{\sigma}^Q$ and $\hat{\sigma}^{Q_{Inv}}$) are estimated by:

$$\hat{\sigma}^Q(j) = \sqrt{\frac{1}{J-j}\sum_{n=1}^{J-j+1} I_{n,j}(Q_{n,j} - \hat{q}(j))^2} \tag{3.4}$$

$$\hat{\sigma}^{Q_{Inv}}(j) = \sqrt{\frac{1}{J-j}\sum_{n=1}^{J-j+1} P_{n,j}(Q_{n,j}^{-1} - \hat{q}_{Inv}(j))^2} \tag{3.5}$$

Slopes of The Regression Lines Without A Constant

To calculate λ^P and λ^I the residuals for F^P, Q^{-1} F^I and Q needs to be calculated first. The residuals for F^P and F^I are calculated by:

$$\mathbf{Res}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j) = \widehat{\mathbf{Res}}(F_{i,j}^P) = \frac{F_{i,j}^P - \hat{f}_j^P}{\hat{\sigma}(\frac{P_{i,j+1}}{P_{i,j}}|\mathbf{P}_j)} = \frac{F_{i,j}^P - \hat{f}_j^P}{\hat{\sigma}^P(j)}\sqrt{P_{i,j}} \tag{3.6}$$

$$\mathbf{Res}(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_j) = \widehat{\mathbf{Res}}(F_{i,j}^I) = \frac{F_{i,j}^I - \hat{f}_j^I}{\hat{\sigma}(\frac{I_{i,j+1}}{I_{i,j}}|\mathbf{I}_j)} = \frac{F_{i,j}^I - \hat{f}_j^I}{\hat{\sigma}^I(j)}\sqrt{I_{i,j}} \tag{3.7}$$

The residuals for Q and Q^{-1} are calculated in a similar way:

$$\mathbf{Res}(\frac{P_{i,j}}{I_{i,j}}|\mathbf{B}_j) = \widehat{\mathbf{Res}}(Q_{i,j}) = \frac{Q_{i,j} - \hat{q}(j)}{\hat{\sigma}(Q_{i,j}|\mathbf{B}_j)} = \frac{Q_{i,j} - \hat{q}(j)}{\hat{\sigma}^Q(j)}\sqrt{I_{i,j}} \tag{3.8}$$

$$\mathbf{Res}(\frac{I_{i,j}}{P_{i,j}}|\mathbf{B}_j) = \widehat{\mathbf{Res}}(Q_{i,j}^{-1}) = \frac{Q_{i,j}^{-1} - \hat{q}_{Inv}(j)}{\hat{\sigma}(Q_{i,j}^{-1}|\mathbf{B}_j)} = \frac{Q_{i,j}^{-1} - \hat{q}_{Inv}(j)}{\hat{\sigma}^{Q_{Inv}}(j)}\sqrt{P_{i,j}} \tag{3.9}$$

The slopes of the regression lines without a constant (λ^I and λ^P) can now be calculated as

$$\widehat{\lambda^P} = \frac{\sum_A \widehat{\mathbf{Res}}(Q_{i,j}^{-1})\widehat{\mathbf{Res}}(F_{i,j}^P)}{\sum_A \widehat{\mathbf{Res}}(Q_{i,j}^{-1})^2} \qquad \text{and} \qquad \widehat{\lambda^I} = \frac{\sum_A \widehat{\mathbf{Res}}(Q_{i,j})\widehat{\mathbf{Res}}(F_{i,j}^I)}{\sum_A \widehat{\mathbf{Res}}(Q_{i,j})^2} \tag{3.10}$$

where A is the set of all (accident years, development years) pairs with four well defined residuals.

3.2 BMCL

3.2.1 Method

BMCL does not change the paid or incurred claims but it re-estimates the parameters (the development factors, the proportionality constants, λ^P and λ^I) in each iteration. BMCL calculates the parameters in the same way as the Liu-Verrall MCL method and all equations and theory in this subsection is from (Liu and Verrall (2010)).

BMCL is done in six steps:

1. Calculate and blockwise resample the four sets of residuals used in the ordinary MCL, multiply them by $\sqrt{\frac{J-j}{J-j-1}}$ and together with the non-bootstrapped conditional expected values $\hat{f}_j^P$, $\hat{f}_j^I$, $\hat{q}$ and $\hat{q}_{Inv}$ create the pseudo-random $\widetilde{F^P}$, $\widetilde{F^I}$, $\widetilde{Q}$ and $\widetilde{Q}^{-1}$.
2. Calculates the new development factors ($\widetilde{f}_j^P$ and $\widetilde{f}_j^I$) and conditional expected values of Q and Q^{-1} ($\widetilde{q}$ and $\widetilde{q}_{Inv}$) with $\widetilde{F^P}$, $\widetilde{F^I}$, $\widetilde{Q}$ and $\widetilde{Q}^{-1}$.
3. Calculate $\widetilde{\lambda^I}$ and $\widetilde{\lambda^P}$ using the pseudo-random residuals.
4. Calculate the proportionality constants ($\widetilde{\sigma}^P$,$\widetilde{\sigma}^I$,$\widetilde{\sigma}^Q$ and $\widetilde{\sigma}^{Q_{Inv}}$) using $\widetilde{F^P}$,$\widetilde{F^I}$,$\widetilde{Q}$ and $\widetilde{Q}^{-1}$ and their conditional expected values $\widetilde{f}_j^P$, $\widetilde{f}_j^I$, $\widetilde{q}$ and $\widetilde{q}_{Inv}$.
5. Calculate $\widetilde{f}_{MCL}^P(i,j)$ and $\widetilde{f}_{MCL}^I(i,j)$ using the bootstrapped parameters (the parameters with a tilde).
6. Estimating future claims using Eq. (2.19) and (2.20) and with future claims calculate the estimated reserves needed.

These six steps are done N number of times which means that an empirical distribution can be created for the reserves.

The Liu-Verrall MCL method has a seventh step, adding a normally distributed error with a zero mean value and a standard deviation linear to the square of the claim to each future claim, however this is not done in BMCL. The reason the step is not used in BMCL, is that it is quite a strong assumption to make that the paid and incurred claims for an accident year are independent, and it can often lead to predictions with Q-values over 1.

3.2.2 Estimating The Parameters

The residuals are calculated in the same way as with the ordinary MCL, see Eq. (3.6) to (3.9). However to correct bootstrap bias the residuals are multiplied by $\sqrt{\frac{J-j}{J-j-1}}$ where J is the total number of development years.

$$\widetilde{\mathbf{Res}} = \widehat{\mathbf{Res}} \cdot \sqrt{\frac{J-j}{J-j-1}} \tag{3.11}$$

Individual Development Factors and Q- and Q_{Inv}-ratios

The bootstrapped individual development factors are created by:

$$\widetilde{F}^P_{i,j} = \frac{\widetilde{\mathbf{Res}}(\widetilde{F}^P_{i,j})\hat{\sigma}^P(j)}{\sqrt{P_{n,j}}} + \hat{f}^P_j \qquad \text{and} \qquad \widetilde{F}^I_{i,j} = \frac{\widetilde{\mathbf{Res}}(\widetilde{F}^I_{i,j})\hat{\sigma}^I(j)}{\sqrt{I_{n,j}}} + \hat{f}^I_j \tag{3.12}$$

In a similar way Q and Q_{Inv} are created by:

$$\widetilde{Q}_{i,j} = \frac{\widetilde{\mathbf{Res}}(\widetilde{Q}_{i,j})\hat{\sigma}^Q(j)}{\sqrt{I_{n,j}}} + \hat{q}(j) \qquad \text{and} \qquad \widetilde{Q}^{-1}_{i,j} = \frac{\widetilde{\mathbf{Res}}(\widetilde{Q}^{-1}_{i,j})\hat{\sigma}^{Q_{Inv}}(j)}{\sqrt{P_{n,j}}} + \hat{q}_{Inv}(j) \tag{3.13}$$

Development Factors and The Expected Value of Q and Q_{Inv}

The bootstrapped development factors ($\widetilde{f}^P_j$ and $\widetilde{f}^I_j$) are created by:

$$\widetilde{f}^P_j = \frac{\sum_{n=1}^{J-j} P_{n,j}\widetilde{F}^P_{i,j}}{\sum_{n=1}^{J-j} P_{n,j}} \qquad \text{and} \qquad \widetilde{f}^I_j = \frac{\sum_{n=1}^{J-j} I_{n,j}\widetilde{F}^I_{i,j}}{\sum_{n=1}^{J-j} I_{n,j}} \tag{3.14}$$

In a similar way the bootstrapped expected value of $\widetilde{Q}$ and $\widetilde{Q}_{Inv}$ ($\widetilde{q}(j)$ and $\widetilde{q}_{Inv}(j)$) are created by:

$$\widetilde{q}(j) = \frac{\sum_{n=1}^{J-j} I_{n,j}\widetilde{Q}_{i,j}}{\sum_{n=1}^{J-j} I_{n,j}} \qquad \text{and} \qquad \widetilde{q}_{Inv}(j) = \frac{\sum_{n=1}^{J-j} P_{n,j}\widetilde{Q}^{-1}_{i,j}}{\sum_{n=1}^{J-j} P_{n,j}} \tag{3.15}$$

Slopes of The Regression Lines Without A Constant

The bootstrapped slopes of the regression lines without a constant ($\widetilde{\lambda}^I$ and $\widetilde{\lambda}^P$) can now be calculated as:

$$\widetilde{\lambda}^P = \frac{\sum_A \widetilde{\mathbf{Res}}(\widetilde{Q}^{-1}_{i,j})\widetilde{\mathbf{Res}}(\widetilde{F}^P_{i,j})}{\sum_A \widetilde{\mathbf{Res}}(\widetilde{Q}^{-1}_{i,j})^2} \qquad \text{and} \qquad \widetilde{\lambda}^I = \frac{\sum_A \widetilde{\mathbf{Res}}(\widetilde{Q}_{i,j})\widetilde{\mathbf{Res}}(\widetilde{F}^I_{i,j})}{\sum_A \widetilde{\mathbf{Res}}(\widetilde{Q}_{i,j})^2} \tag{3.16}$$

where A is the set of all (accident years, development years) pairs with four well defined residuals.

Proportionality Constants

The proportionality constants for paid and incurred claims ($\widetilde{\sigma}^p$ and $\widetilde{\sigma}^I$) are estimated by:

$$\widetilde{\sigma}^P(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} P_{n,j}(\widetilde{F}^P_{i,j} - \widetilde{f}^P_j)^2} \tag{3.17}$$

$$\widetilde{\sigma}^I(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} I_{n,j}(\widetilde{F}^I_{i,j} - \widetilde{f}^I_j)^2} \tag{3.18}$$

In the same way the proportionality constants for Q-ratio and Q_{Inv}-ratio ($\widetilde{\sigma}^Q$ and $\widetilde{\sigma}^{Q_{Inv}}$) are estimated by:

$$\widetilde{\sigma}^Q(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} I_{n,j}(\widetilde{Q}_{n,j} - \widetilde{q}(j))^2} \tag{3.19}$$

$$\widetilde{\sigma}^{Q_{Inv}}(j) = \sqrt{\frac{1}{J-j-1}\sum_{n=1}^{J-j} P_{n,j}(\widetilde{Q}^{-1}_{n,j} - \widetilde{q}_{Inv}(j))^2} \tag{3.20}$$

$\widetilde{\sigma}^Q(j)$ and $\widetilde{\sigma}^{Q_{Inv}}(j)$ both have a factor $\sqrt{\frac{1}{J-j-1}}$ instead of $\sqrt{\frac{1}{J-j}}$ as done in the ordinary MCL, this is because the blockwise bootstrapping does only resample the (accident year, development year) pair which has all four residuals. This means that there is one less residual resampled per development year than there are Q and Q_{Inv} residuals.

3.2.3 One Year Reserve Risk BMCL

Calculating r_1 using BMCL is done by the following steps:

1. Create $\mathbf{D}_n(J+1)$ by estimating the paid and incurred claims at time $J+1$ with BMCL.
2. Input $\mathbf{D}_n(J+1)$ into MCL to estimate $C^{\mathbf{K}}_{i,J}|\mathbf{D}_n(J+1)$.
3. Calculate $\epsilon(n)$.

Repeat step 1 to 3 N number of times, and then

4. Create an empirical PDF of $\epsilon(n)$.
5. Use the empirical PDF to calculate r_1 using Eq. (2.28).

As empirical PDFs converge slower at the tails to the true PDFs, a high N is needed to get a good estimate of the 99.5 percentile of ϵ and thereby r_1.

3.3 Data problem

As mentioned in the beginning of this chapter MCL and BMCL assumes that paid and incurred data can be written as complete triangles and with cumulative claims ≥ 0 for each element. However this is often not the case. In the following sections several types of data discrepancies will be examined and suggestions on how to deal with the problems that arises from these data discrepancies will be proposed. In general there are three ways of dealing with a data discrepancy:

1. Adapting MCL and BMCL to be able to handle the discrepancy.
2. Change or estimate the data in order to "repair" it.
3. Skip the entire development or accident year.

The first option is preferred as it keeps all the good data and does not add a risk of inducing errors from the changing of data. However in some cases the first option cannot be done while maintaining a good prediction model and one has to consider the second or third option.

A dataset can have several types of problems at the same time and in such cases a decision will have to be made on a case by case basis on how it should be handled.

3.3.1 Missing Data

The missing data can be put into four different groups, types 1 to 4. Type 1 and 2 are missing data points on accident years in which there is no known data in the preceding development years, while Type 3 and 4 there is known data in the preceding development years. Type 1 and 3 are symmetrical, meaning that an (accident year, development year) pair have both unknown paid and incurred data, while Type 2 and 4 are asymmetrical, meaning that an (accident year, development year) pair only have paid or incurred unknown. See Table 3.2 for a dataset with all four types of missing data.

	1	2	3	4	$J=5$
1	NA	$P_{1,2}$	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	$P_{2,1}$	NA	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	NA			
$J=5$	$P_{5,1}$				

	1	2	3	4	$J=5$
1	NA	NA	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	$I_{2,1}$	NA	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	$I_{4,2}$			
$J=5$	$I_{5,1}$				

Table 3.2: Dataset with all four types of missing data.{1,1} is Type 1 missing data, {1,2} is Type 2 missing data, {2,2} is Type 3 missing data and {4,2} is Type 4 missing data.

Type 1 Missing Data

With Type 1 missing data MCL and BMCL are adapted to ignore the missing data points and only use data from (accident years, development years) pairs with paid and incurred data when calculating the parameters. This way no data needs to be added or changed and all of the available data can be used, which means that no information is lost and no error can be induced by data manipulation. For this reason this is the chosen way to deal with Type 1 missing data.

Another way of dealing with Type 1 missing data is to approximate the missing data by CLM. This way f^I and f^P stays the same, but q, q_{Inv}, λ^P and λ^I do not stay the same and an possible error has then been induced. Therefore this method is not used in this thesis.

Nearly all of the missing data points in the datasets examined are Type 1 missing data.

	1	2	3	4	$J=5$
1	NA	NA	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	NA	$P_{2,2}$	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	$P_{4,2}$			
$J=5$	$P_{5,1}$				

	1	2	3	4	$J=5$
1	NA	NA	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	NA	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	$I_{4,2}$			
$J=5$	$I_{5,1}$				

Table 3.3: Dataset with Type 1 missing data

Example

With the suggested method and the data in Table 3.3 $\hat{f}_1^P$ and $\hat{f}_2^P$ are approximated by:

$$\hat{f}_1^P = \frac{\sum_{n=3}^{4} P_{n,j+1}}{\sum_{n=3}^{4} P_{n,j}} \quad \text{and} \quad \hat{f}_2^P = \frac{\sum_{n=2}^{3} P_{n,j+1}}{\sum_{n=2}^{3} P_{n,j}} \tag{3.21}$$

In these equations the sums do not include $P_{1,1}, P_{2,1}$ and $P_{1,2}$ as the data is missing. By excluding the missing data points f^P is well defined for $1 \ldots J-1$ again. The same is done for f^I, q and q_{Inv}. In a similar way $\hat{\sigma}^P(1)$ and $\hat{\sigma}^P(2)$ are approximated by:

$$\hat{\sigma}^P(1) = \sqrt{\frac{1}{1}\sum_{n=3}^{4} P_{n,1}(\frac{P_{n,2}}{P_{n,1}} - \hat{f}_1^P)^2} \quad \text{and} \quad \hat{\sigma}^P(2) = \sqrt{\frac{1}{1}\sum_{n=2}^{3} P_{n,2}(\frac{P_{n,3}}{P_{n,2}} - \hat{f}_2^P)^2} \tag{3.22}$$

By changing P to I one approximates $\hat{\sigma}^I(1)$ and $\hat{\sigma}^I(2)$.

$\hat{\sigma}^Q(1)$ and $\hat{\sigma}^Q(2)$ are approximated by:

$$\hat{\sigma}^Q(1) = \sqrt{\frac{1}{2}\sum_{n=3}^{5} I_{n,1}(\frac{P_{n,1}}{I_{n,1}} - \hat{q}(1))^2} \quad \text{and} \quad \hat{\sigma}^Q(2) = \sqrt{\frac{1}{2}\sum_{n=2}^{4} I_{n,2}(\frac{P_{n,2}}{I_{n,2}} - \hat{q}(2))^2} \tag{3.23}$$

$\sigma^{Q_{Inv}}(1)$ and $\sigma^{Q_{Inv}}(2)$ are approximated by switching P and I.

The residuals are calculated the same way as with complete data, and the slopes of regression with no constant are approximated by Eq. (3.9), the same as with the complete data, only with a redefined A that no longer contains the elements $\{1,1\}, \{2,1\}$ and $\{1,2\}$.

Type 2 Missing Data

The suggested way, and the one chosen in this thesis, of handling Type 2 missing data is to treat all (accident years, development years) pairs as completely missing data if either the paid or incurred data point is missing. Another way is to approximate the missing paid or incurred data with CLM for each (accident years, development years) pair which only have the paid or incurred data. This way all information is used, but error can be induced by the approximation of the missing data.

	1	2	3	4	$J=5$
1	NA	$P_{1,2}$	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	NA	$P_{2,2}$	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	$P_{4,2}$			
$J=5$	$P_{5,1}$				

	1	2	3	4	$J=5$
1	NA	NA	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	$I_{2,1}$	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	$I_{4,2}$			
$J=5$	$I_{5,1}$				

Table 3.4: Dataset with Type 1 and Type 2 missing data

Example

With the suggested method and data in Table 3.4 $P_{1,2}$ and $I_{2,1}$ are treated as missing data, which means that Table 3.3 and Table 3.4 give the exact same results independently of the values of $P_{1,2}$ and $I_{2,1}$. This is a weakness of the method.

Using the other method, to estimate the missing data by CLM, does not give identical results independent of the values of $P_{1,2}$ and $I_{2,1}$, which is reasonable. However as always there is the risk of inducing errors when estimating data.

Type 3 Missing Data

Type 3 missing data is handled using the same method that is used for Type 1 missing data, as long as the unknown claims are not in the latest known calendar year. For Type 3 missing data in the latest known calendar year the unknown claims needs to be approximated to be able to calculate future claims. The suggested way of doing this approximation is by the MCL method.

Type 4 Missing Data

With type 4 missing data one can get a better estimation of the missing data than with type 2 using MCL. With type 4 missing data one should decide on a case by case basis which method to use. None of the 38 datasets considered in this thesis had any Type 4 missing data.

3.3.2 Problematic Data

Zero Values

The zero value is used to approximate q, q_{Inv} and f^P or f^I depending on if the zero value is a paid or an incurred claim. The zero value claim is not used to calculate Q, Q^{-1},F^I nor F^P as when a paid claim is zero then Q^{-1} and F^P are not defined and when an incurred claim is zero then Q and F^I are not defined. This means that one does not calculate the residual for the (accident year, development year) pair with the zero value claim and it is not used in calculating λ^P and λ^I.

If the zero value claim is in the final calendar year the f_{MCL} for that (accident year, development year) pair is not defined, which means that the future claims will be undefined. This problem is solved by rewriting Eq. (2.11) and (2.1) into Eq. (3.24). The same can be done for incurred claims by rewriting Eq. (2.12) and (2.1) into Eq. (3.25).

$$P_{i,j+1} = P_{i,j} f^P_{MCL}(i,j) = P_{i,j+1}(f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j}(Q^{-1}_{i,j} - q_{Inv}(j)))$$

$$= P_{i,j} f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j}(Q^{-1}_{i,j} P_{i,j} - q_{Inv}(j) P_{i,j}) = P_{i,j} f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j}(\frac{I_{i,j}}{P_{i,j}} P_{i,j} - q_{Inv}(j) P_{i,j}) =$$

$$= P_{i,j} f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j}(I_{i,j} - q_{Inv}(j) P_{i,j})$$

$$P_{i,j+1} = P_{i,j} f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j}(I_{i,j} - q_{Inv}(j) P_{i,j}) \tag{3.24}$$

$$I_{i,j+1} = I_{i,j} f^I_j + \lambda^I \cdot \frac{\sigma^I_j}{\sigma^Q_j}(P_{i,j} - q(j) I_{i,j}) \tag{3.25}$$

	1	2	3	4	$J=5$
1	0	$P_{1,2}$	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	$P_{2,1}$	$P_{2,2}$	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	$P_{4,2}$			
$J=5$	0				

	1	2	3	4	$J=5$
1	$I_{1,1}$	$I_{1,2}$	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	0	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	0			
$J=5$	$I_{5,1}$				

Table 3.5: Dataset with zero value claims.

Negative Values

Neither MCL nor BMCL is designed to handle negative values. In MCL and BMCL both the Q and Q_{Inv} will be negative if a paid or incurred claim is negative, meaning that the method will try to correct it for both a too high Q-ratio and a too high Q_{Inv}-ratio at the same time, and also the approximation of λ^P and λ^I will use these misrepresenting Q and Q_{Inv} values.

Negative value claims will not be used and will either have to be assumed to be missing data, or manipulated, or the entire accident year not used at all, depending on the dataset.

	1	2	3	4	$J=5$
1	$P_{1,1}<0$	$P_{1,2}$	$P_{1,3}$	$P_{1,4}$	$P_{1,5}$
2	$P_{2,1}$	$P_{2,2}$	$P_{2,3}$	$P_{2,4}$	
3	$P_{3,1}$	$P_{3,2}$	$P_{3,3}$		
4	$P_{4,1}$	$P_{4,2}$			
$J=5$	$P_{5,1}$				

	1	2	3	4	$J=5$
1	$I_{1,1}$	$I_{1,2}$	$I_{1,3}$	$I_{1,4}$	$I_{1,5}$
2	$I_{2,1}$	$I_{2,2}$	$I_{2,3}$	$I_{2,4}$	
3	$I_{3,1}$	$I_{3,2}$	$I_{3,3}$		
4	$I_{4,1}$	$I_{4,2}$			
$J=5$	$I_{5,1}$				

Table 3.6: Dataset with negative value claims

3.4 Bad Parametric Values

3.4.1 Negative λ^P or λ^I

When $\lambda^P < 0$ or $\lambda^I < 0$ MCL and BMCL can have some problems. The advantage of MCL is that it converges Q-values as time progresses by increasing the paid and decreasing the incurred development factors for accident years with lower than average Q-values, and vice versa. This is done by Eq. (2.13) and (2.14), however if λ^P or λ^I is negative, this gets reversed.

$$f^P_{MCL}(i,j) = f^P_j + \Delta f^P_{i,j}, \qquad \Delta f^P_{i,j} = \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j} \cdot (Q^{-1}_{i,j} - q_{inv}(j)) \tag{2.13}$$

$$f^I_{MCL}(i,j) = f^I_j + \Delta f^I_{i,j}, \qquad \Delta f^I_{i,j} = \lambda^I \cdot \frac{\sigma^I_j}{\sigma^Q_j} \cdot (Q_{i,j} - q(j)) \tag{2.14}$$

For an accident year with a lower than average Q-value, $(Q^{-1}_{i,j} - q_{inv}(j)) < 0$, in a dataset with $\lambda^P \leq 0$, the individual paid development factor is calculated as:

$$f^P_{MCL}(i,j) = f^P_j + \lambda^P \cdot \frac{\sigma^P_j}{\sigma^{Q_{Inv}}_j} \cdot (Q^{-1}_{i,j} - q_{inv}(j)) \rightarrow f^P_{MCL}(i,j) = f^P_j - D \rightarrow f^P_{MCL}(i,j) < f^P_j \tag{3.26}$$

where D is a positive constant.

Eq. (3.26) gives that for accident years which has a lower than average Q-value the paid individual development factors are lower than with CLM. This decreases the already too low Q-value the following development year, see Fig. 3.1

If both of the λs are negative then they will both diverge the Q-values and the entire reason to use MCL is not only lost but it actually makes the problem it was supposed to solve worse. If one λ is negative and one is positive then one of them will try to diverge the Q-values and one to converge them, meaning that in some cases MCL will work and make the Q-values converge and in some cases it will not. Looking at Fig. 3.2, which has a negative λ^I, it can be seen that converging Q-values is possible despite a negative λ.

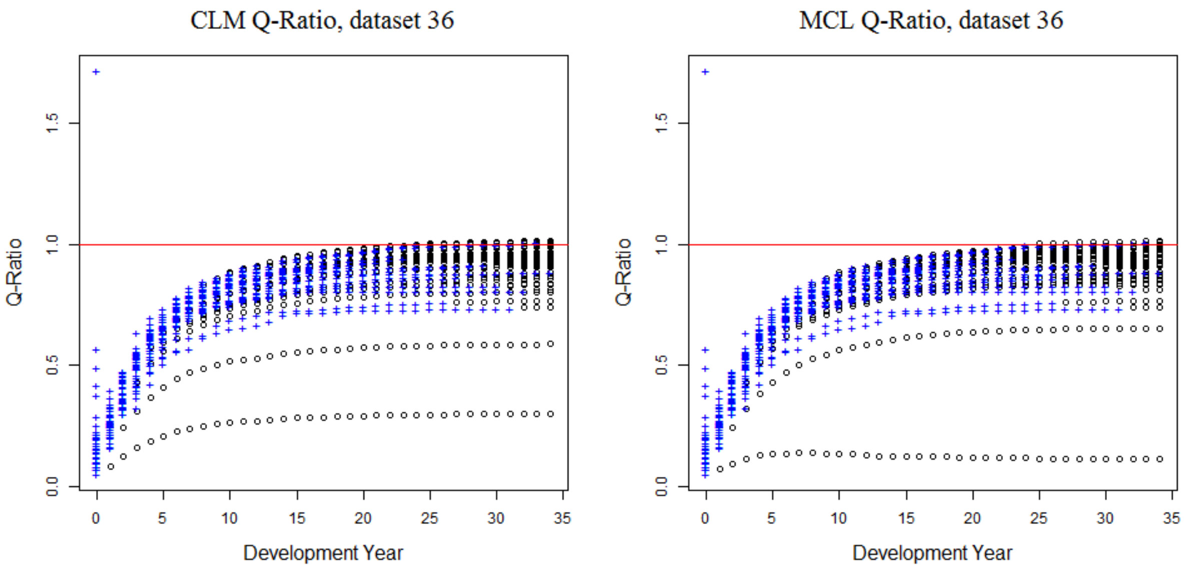

Figure 3.1: Dataset 36, negative λ^P, blue "+" are real claim data, black "o" is predicted claim data. The left is predicted by CLM and the right figure by MCL. In the left figure the Q-value convergence near 1, that one expects from MCL, does not exist. Looking at the accident year with the lowest Q-value one can see that MCL makes it diverge from the average Q instead of converge towards it.

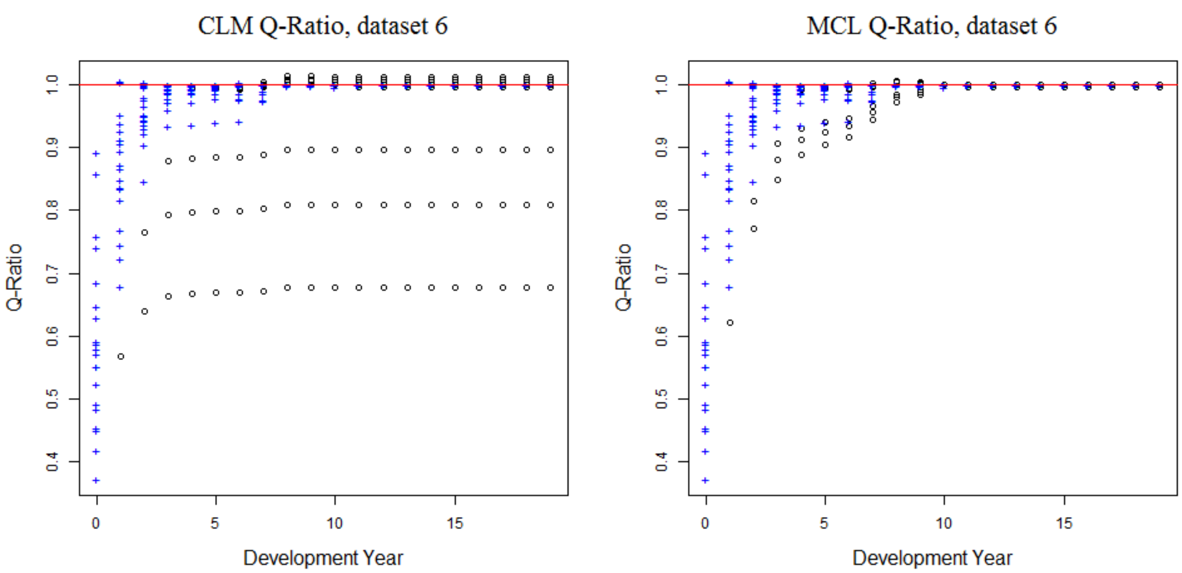

Figure 3.2: Dataset 6, negative λ^I, blue "+" are real claim data, black "o" is predicted claim data. The right figure is predicted by MCL and the left by CLM. The predictions by MCL converge to the Q-value of 1 despite the negative λ^I.

3.4.2 High Sigma-ratio

Eq. (2.13) and (2.14) shows that the individual development factors in MCL and BMCL, $f_{MCL}(i,j)$, can be written as the CLM development factor, f_j, plus a correction term, $\Delta f_{i,j}$, that is accident year dependent. The correction term, as long as λ^P and λ^I are positive, increases the Q-value of the next development year for accident years that have a lower than average Q-value and vice versa.

The correction terms are linearly dependent of the "Sigma-ratios" ,$\frac{\sigma_j^P}{\sigma_j^{QInv}}$ and $\frac{\sigma_j^I}{\sigma_j^Q}$. The Sigma-ratios have for the most part a value below 1, often even below 0.1. However the Sigma-ratio can be much larger, in a few cases up to 20. In these cases the correction from $\Delta f_{i,j}$ can become far too large, leading to accident years with lower than average Q-values to get a far to high Q-value next development year, see the right graph in Fig. 3.3.

To help alleviate this problem an upper limit can be put on the Sigma-ratio. This improves the methods ability to produce good predictions without changing the correlations, see Fig. 3.3 for a comparison of predictions made by MCL with and without an upper limit on the Sigma-ratio.

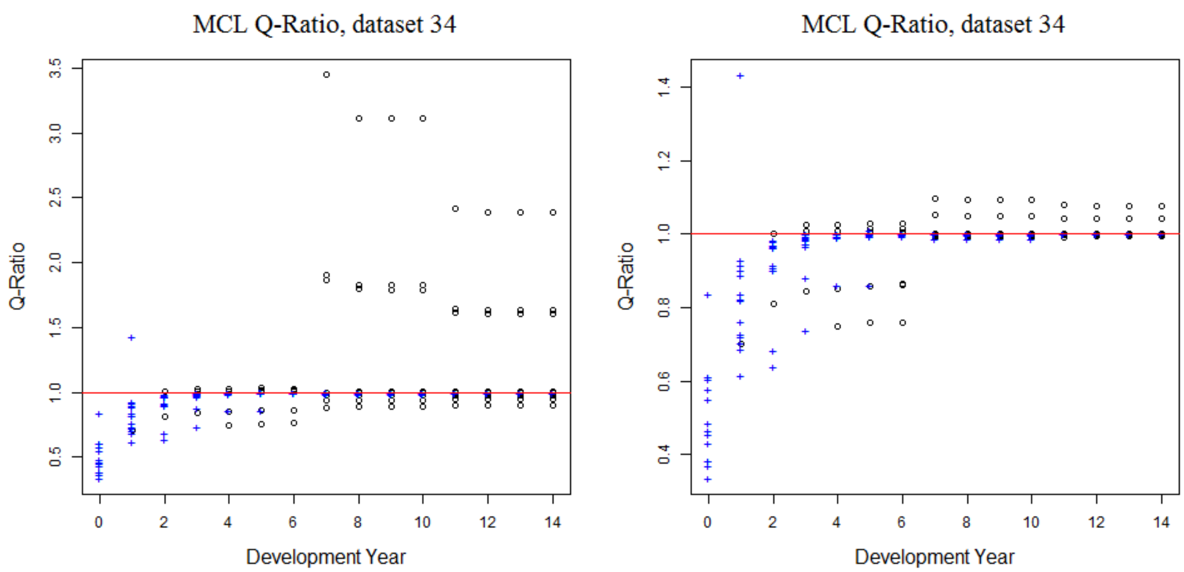

Figure 3.3: Both figures are predictions of the same dataset predicted with MCL. The left has no upper limit for the Sigma-ratio while the right one has a upper limit of 5. As one can see at year seven MCL gives extreme Q-values for the unlimited case, and while the limited case get Q-values greater than 1 they are not nearly as extreme.

Chapter 4

Result and Analysis

4.1 The Datasets

The data used to calculate the results in this thesis are from 36 different portfolios from both Trygg-Hansa and Codan, and from different line of businesses. The insurance portfolios are differing in number of development years, in size of claims and in number of unknown data points. These datasets will simply be referred to as "dataset 1" to "dataset 36". Any missing or strange data will be dealt with as written in section 3.3.

Dataset A and B will also be examined but they will not be in the results when considering dataset 1 to 36.

4.2 MCL

In this section the result of MCL will be discussed, the result from BMCL will be in the next section.

4.2.1 Q-ratios

In 37 of 38 different datasets (A,B and 1-35) the Q-values can be seen to converge better with MCL than with CLM, however the difference in convergence vary greatly from dataset to dataset. For example, for dataset A and B, see Fig. 4.1, the Q-values converge better to 1 when using MCL, however the size of the improvement is much greater for dataset A than for dataset B.

The one dataset that does not get helped by MCL to converge the Q-values is dataset 36, see Fig. 3.1 and the reasons behind this is discussed in subsection 3.4.1.

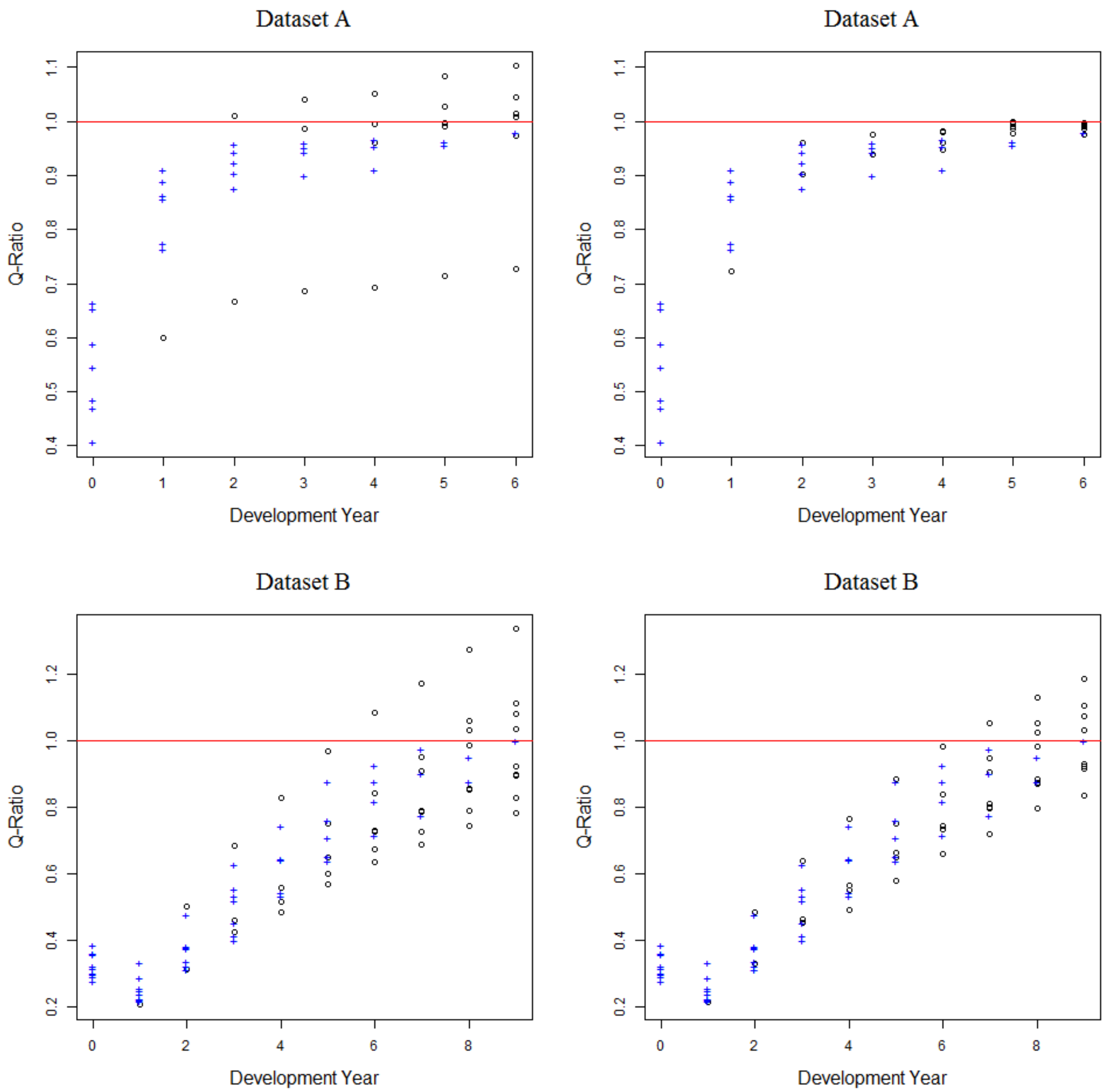

Figure 4.1: Dataset A on the top and dataset B on the bottom, prediction by CLM on the left, MCL to the right. Development year on the X-axis and the P/I-ratio on the Y-axis. The red line is at theP/I-ratio equal to 1.

	1	2	3	4	$J=5$
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$	$C_{1,5}$
1	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$	$C_{2,5}$
1	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,4}$	$C_{3,5}$
1	$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$	$C_{4,5}$
$J=5$	$C_{5,1}$	$C_{5,2}$	$C_{5,3}$	$C_{5,4}$	$C_{5,5}$

Table 4.1: $J = 5$. Cumulative claim triangle with the accident years as rows and the development years as columns. $C_{i,j}$ is the cumulative paid or incurred claim for accident year i after j development years. The calendar year is $i + j - 1$.

4.2.2 Reserve and Future claims

In this section the error of estimated reserves and future claims from MCL and CLM are compared. By using the filled triangles, where all of the accident years have J known development years, see Table 4.1, and Eq. (2.23) one can calculate an ultimate reserve, R_∞^{Filled}.

By considering the same triangle, but non-filled, and calculating estimated ultimate reserves by MCL and CLM (R_∞^{MCL} and R_∞^{CLM}) the differences between the estimated and the real reserves can then be calculated. The difference between the estimated and real reserves, ΔR, are defined as:

$$\Delta R_{MCL} = R_\infty^{Filled} - R_\infty^{MCL} \qquad \Delta R_{CLM} = R_\infty^{Filled} - R_\infty^{CLM} \tag{4.1}$$

The claims at calendar year $J + 1$, $C_{i,J-i+2}^{\mathbf{K},MCL}$ and $C_{i,J-i+2}^{\mathbf{K},CLM}$ $\forall i$, will also be compared to the real claims, $C_{i,J-i+2}$. The comparison of claims at calendar year $J + 1$ will be done by comparing ΔC_{MCL} and ΔC_{CLM}:

$$\Delta C_{MCL} = \sum_{i=2}^{J}(C_{i,J-i+2}^{\mathbf{K},MCL} - C_{i,J-i+2}^{\mathbf{K},Real}) \qquad \Delta C_{CLM} = \sum_{i=2}^{J}(C_{i,J-i+2}^{\mathbf{K},CLM} - C_{i,J-i+2}^{\mathbf{K},Real}) \tag{4.2}$$

By comparing the absolute ΔR_{MCL}, ΔR_{CLM}, ΔC_{MCL} and ΔC_{CLM} one can evaluate the accuracy of the predictions of MCL compared to CLM.

In this thesis MCL is examined using several different sets of datasets:

1. All of the 36 datasets except dataset 21 and 35.
2. The datasets in which MCL have positive λ^I and λ^P.
3. The datasets in which MCL have positive λ^I and λ^P and no Q-value above 1.02.

The comparison of MCL and CLM will also be done with all available accident years and then only with accident years after 1999.

All Accident Years, All Datasets

When considering all insurance portfolios with all accident years, Table 4.2, MCL was better than CLM with paid claims in predicting both the ultimate reserve and the paid claims one year in the future in a majority of the datasets. MCL was also slightly better than CLM with incurred data in predicting the incurred claims one year in the future, however it was worse in predicting the ultimate reserve.

n=34	Incurred	Paid
Number of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	15	19
% of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	44.1%	55.9%
Number of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	18	20
% of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	52.9%	58.8%

Table 4.2: Dataset 1-36 except dataset 21 and dataset 35.

All Accident Years, All Datasets with $\lambda^I \geq 0$ and $\lambda^P \geq 0$

By excluding the datasets with negative λ^I or λ^P one could assume that MCL will perform better. MCL is superior to CLM with paid data in predicting both the ultimate reserve and claims one year in the future, however it was slightly worse than CLM with incurred data at predicting both, see Table 4.3.

n=25	Incurred	Paid
Number of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	12	16
% of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	48%	64%
Number of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	12	17
% of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	48%	68%

Table 4.3: Only the datasets with $\lambda^I \geq 0$ and $\lambda^P \geq 0$.

All Accident Years, All Datasets with $\lambda^I \geq 0$, $\lambda^P \geq 0$ and $Q_{i,j} \leq 1.02, \quad \forall(i,j)$

By further excluding datasets that do not have $Q_{i,j} \leq 1.02, \forall(i,j)$ one removes datasets which do not have the Q-pattern which MCL is trying to emulate. This should be an advantage for MCL. With this small set of datasets, MCL outperforms CLM with both paid and incurred data in predicting both the ultimate reserve and the claims one year in the future. However drawing conclusions one have to keep in mind there are only 14 datasets compared.

n=14	Incurred	Paid
Number of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	8	9
% of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	57.1%	64.3%
Number of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	8	8
% of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	57.1%	57.1%

Table 4.4: Only the datasets with $\lambda^I \geq 0$, $\lambda^P \geq 0$ and no Q-Values greater than 1.02.

Accident Years after 1999, All Datasets

When considering all accident years, for some of the datasets the first being as early as 1976, the final one being in 2013 one makes two implicit assumptions that:

1. It makes no significant difference in how a claim from 1976 and a claim from 2013 develop over time.

2. The incurred claims are reported and estimated in the same way from 1976 until 2013.

The first one can be discussed, and the second one is not true. As MCL is a more complex method than CLM any induced error from incorrect or incomplete data is likely to induce larger errors into BMCLs predictions.

Comparing Table 4.2 and Table 4.5 one can see that MCL improves its predictive abilities compared to CLM with paid data when only considering accident years after 1999. The percentage of superior predictions of MCL increases significantly, from 55.9% to 65.7% and 58.8% to 68.6% in ultimate reserve and claims one year in the future respectively.

n=35	Incurred	Paid
Number of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	17	23
% of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	48.6%	65.7%
Number of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	18	24
% of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	51.4%	68.6%

Table 4.5: Datasets 1 to 35, considering only accident years after 1999.

Accident Years after 1999, All Datasets with $\lambda^I \geq 0$, $\lambda^P \geq 0$

Sorting for datasets with $\lambda^I \geq 0$ and $\lambda^P \geq 0$ when only using accident years after 1999 does not have the expected outcome. The relative accuracy of prediction for MCL decreases instead of increases, as in the case with all accident years. No mathematical or theoretical reason why the relative accuracy of prediction should fall can easily be seen. It is possible it is only stochastic chance as there are only 24 datasets, however it should be noted and future research can examine this strange behavior with different datasets.

n=24	Incurred	Paid
Number of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	10	15
% of $\lvert\Delta R_{MCL}\rvert < \lvert\Delta R_{CLM}\rvert$	41.7%	62.5%
Number of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	11	16
% of $\lvert\Delta C_{MCL}\rvert < \lvert\Delta C_{CLM}\rvert$	45.8%	66.7%

Table 4.6: The datasets that when considering only accident years after 1999 have $\lambda^I \geq 0$ and $\lambda^P \geq 0$.

4.3 BMCL

In this section the BMCL will be examined. The things that will be examined are:

1. How the convergence of mean and VaR depend on the number of simulations.
2. How the mean value of the BMCL reserve and MCL reserve compares.
3. The estimated future distribution of the ultimate reserve.
4. The OCDR and the estimated one year reserve risk.

4.3.1 Number of Simulations and Convergence

When using stochastic simulation models there is always a simulation error induced by using a finite number of simulations. In order to see the approximate size of the simulation error the BMCL program is run five times with 10,000 simulations each. By checking the mean and $VaR_{0.995}$ of subsets of the 10,000 simulations each time the program run one can see how well they converge when increasing the size of the subsets.

Fig. 4.2 and 4.3 clearly shows that the means and $VaR_{0.995}$ converge as sample size increases, however at 4,000 to 5,000 samples the rate of convergence seems to slow down. As more simulations will always decrease the simulation error and one does not have infinite computing power and memory one has to make a decision where the decrease in simulation error is not worth the extra simulations. Due to these constrains the 36 different dataset will only be simulated 4,000 times for R_∞ and for r_1.

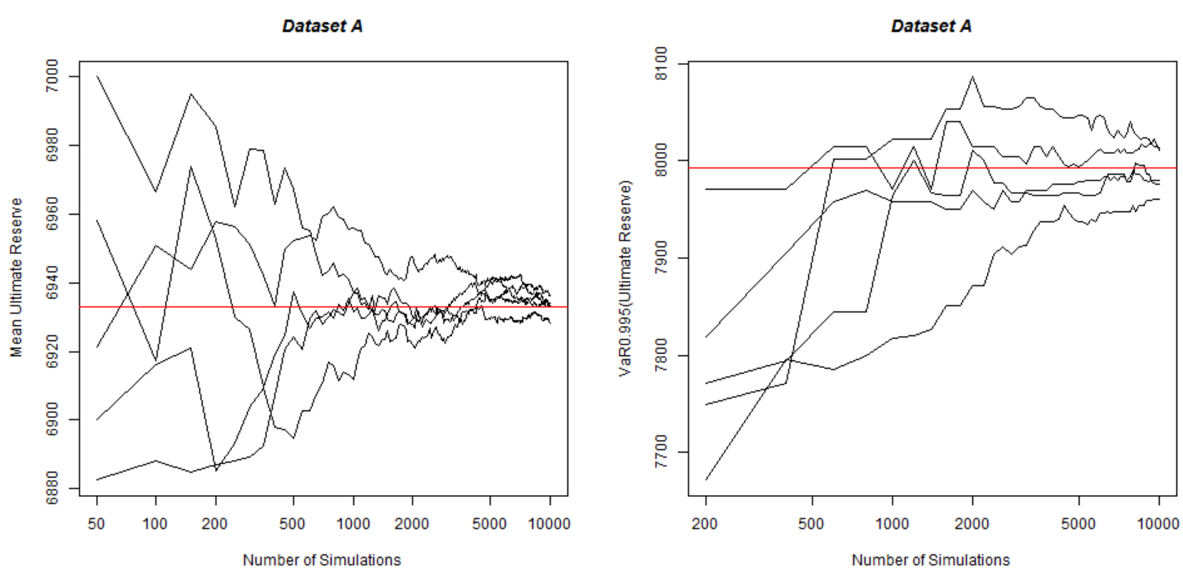

Figure 4.2: Convergence of mean and $VaR_{0.995}$. Mean values in the left graph and $VaR_{0.995}$ in the right one.

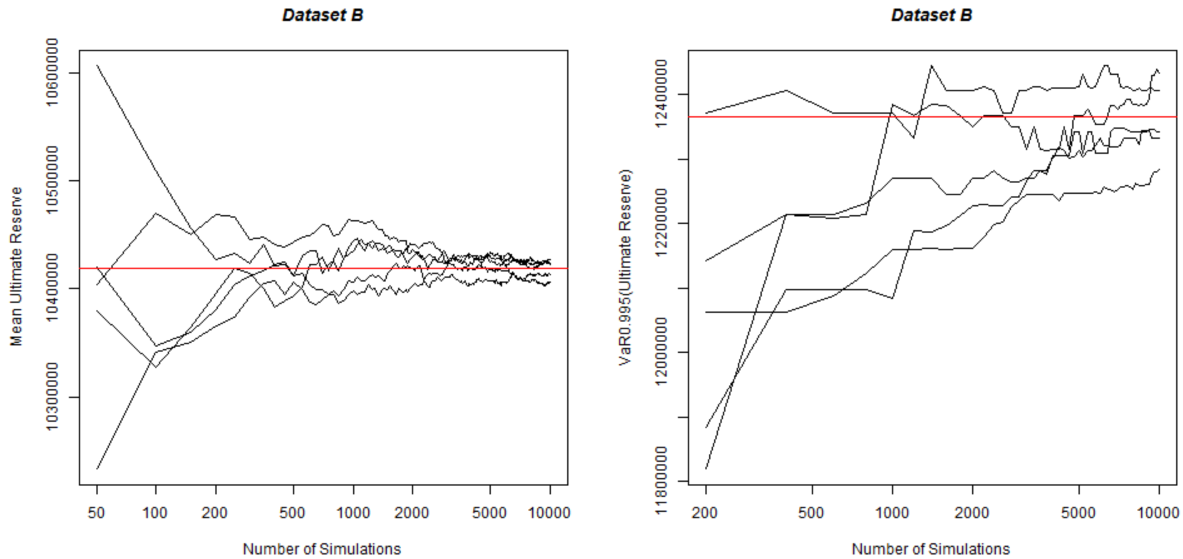

Figure 4.3: Convergence of mean and $VaR_{0.995}$. Mean values in the left graph and $VaR_{0.995}$ in the right one.

4.3.2 Mean BMCL Reserve and MCL Reserve

The mean of the ultimate reserves estimated by BMCL does not seem to converge to ultimate reserve estimated by MCL. The difference in value between the MCL ultimate reserve and the mean BMCL ultimate reserve is highly dependent on the dataset. By examining the relative difference ΔB, see Eq. (4.3), it can be seen that additional simulations have some effect on ΔB, however it would seem like ΔB is not solely responsible for $\Delta B > 0$ (see Table 4.7).

$$\Delta B = \frac{|R_\infty^{MCL} - E[\bar{R}_\infty^{BMCL}]|}{R_\infty^{MCL}} \tag{4.3}$$

The 12.5 % of reserve predictions (9 out of 72) that had a $\Delta B > 0.2$ for 4,000 simulations all had problems with the data or with parametric values. Each of these datasets had missing data, Q-values over 1.02, negative λ and/or high Sigma-ratio. However the datasets which have reserve predictions with $0.05 < \Delta B < 0.2$ does not all have these problems. It would seem that there are datasets that do not have any of the previous data or parametric problems with $R_\infty^{MCL} \neq E[\bar{R}_\infty^{BMCL}]$ even when the number of simulations goes to infinity.

n=36	$\Delta B < 0.02$	$0.02 < \Delta B < 0.05$	$0.05 < \Delta B < 0.2$	$\Delta B > 0.2$
1000 simulations	41.7%	22.2%	27.8%	8.3%
4000 simulations	51.4%	13.9%	22.2%	12.5%

Table 4.7: Dataset A, B and 1 to 36 except dataset 21 and 35.

To make the distributions given by BMCL have a mean of R_∞^{MCL} the reserves estimated by BMCL are multiplied by a factor:

$$R_\infty^{BMCL}(n) = \frac{R_\infty^{MCL}}{E[\bar{R}_\infty^{BMCL}]}\bar{R}_\infty^{BMCL}(n) \tag{4.4}$$

4.3.3 Distribution of Ultimate Reserve and Ultimate Reserve Risk

The distributions of ultimate reserves are made by BMCL with 4,000 simulations. Fig. 4.4 and 4.5 show that for Dataset B the variance differ greatly between the ultimate reserve distributions from the paid and incurred claims and the mean ultimate reserve differ for dataset A and dataset 4. However for dataset 1, 2 and 3 the estimated distributions are very close for the paid and incurred claims. To easily compare r_∞ for different datasets a new variable is defined as:

$$\delta r_\infty = \frac{2r_\infty}{R_\infty^{MCL,Paid} + R_\infty^{MCL,Incurred}}$$

	Dataset A	Dataset B	Dataset 1	Dataset 2	Dataset 3	Dataset 4
δr_∞^{Paid}	0.148	0.189	0.618	0.646	0.471	0.0782
$\delta r_\infty^{Incurred}$	0.161	0.0444	0.640	0.632	0.613	0.0971

Table 4.8: δr_∞ for dataset A, B and 1 to 4.

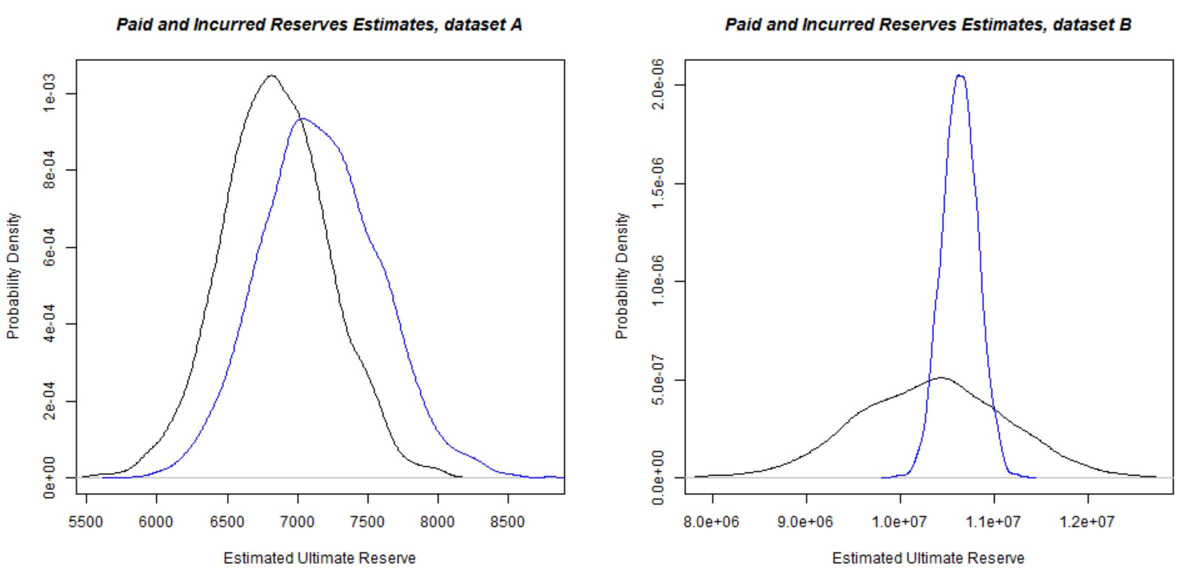

Figure 4.4: Estimated ultimate reserve distribution, blue is from the incurred triangle and black is from the paid triangle.

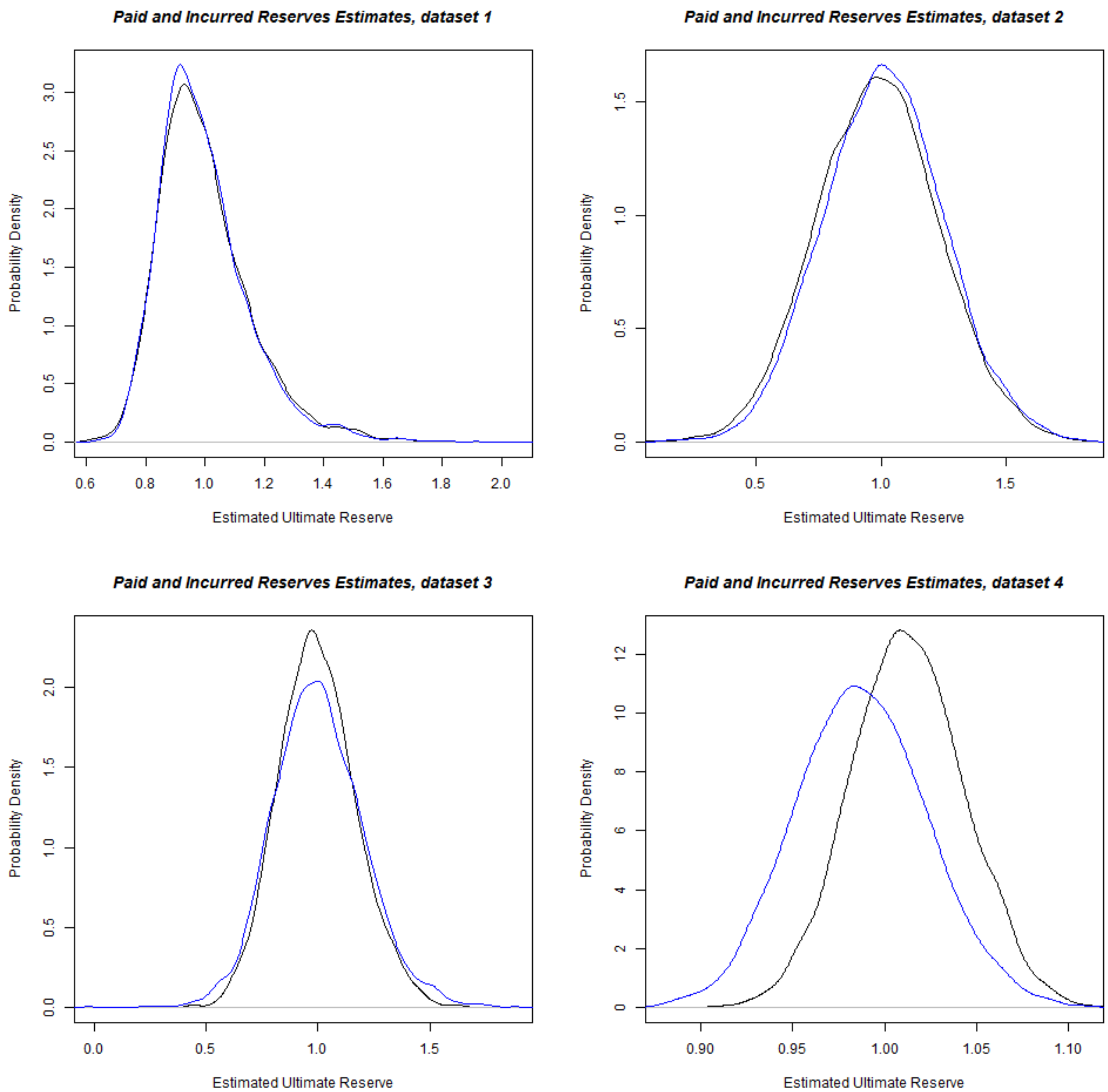

Figure 4.5: Estimated ultimate reserve distribution, blue is from the incurred triangle and black is from the paid triangle. The unit on the X-axis is $0.5R_{\infty}^{MCL,Paid} + 0.5R_{\infty}^{MCL,Incurred}$

.

4.3.4 OCDR and The One Year Reserve Risk

When estimating the distribution of OCDR and the one year reserve risk, r_1, there is a problem. The mean of ϵ calculated with BMCL is not equal to zero, nor does it seem to converge to zero when the number of simulations increases as expected (Lysenko, Merz and Wüthrich (2009)). To work around this problem and get a reasonable estimate of r_1 one makes the approximation of:

$$E_{B(J+1)}[E[C_{i,J}^{\mathbf{K}}|\mathbf{B}_n(J+1)]|\mathbf{B}(J)] = C_{i,J}^{\mathbf{K}} \tag{4.5}$$

Eq. (2.27) and (4.5) gives

$$\epsilon(n) = \sum_{i=2}^{J}(C_{i,J}^{\mathbf{K}}|\mathbf{B}_n(J+1) - E_{B(J+1)}[E[C_{i,J}^{\mathbf{K}}|\mathbf{B}_n(J+1)]|\mathbf{B}(J)]) \tag{4.6}$$

With Eq. (4.6) the PDF of OCDR is centered around zero and r_1 can be calculated. In the same way as in subsection 4.2.3 a new variable is defined as:

$$\delta r_1 = \frac{2r_1}{R_\infty^{MCL,Paid} + R_\infty^{MCL,Incurred}}$$

Comparing Fig. 4.4 and 4.5 to Fig. 4.6 and 4.7 it can be seen that OCDR has a lower variance than the ultimate reserves BMCL estimate. This is expected as the OCDR only includes risks occurring in the next calendar year while the ultimate reserve distributions looks at the entire claim development time. It is also shown that the difference in variance between the ultimate reserve and the OCDR is highly dependent on the dataset.

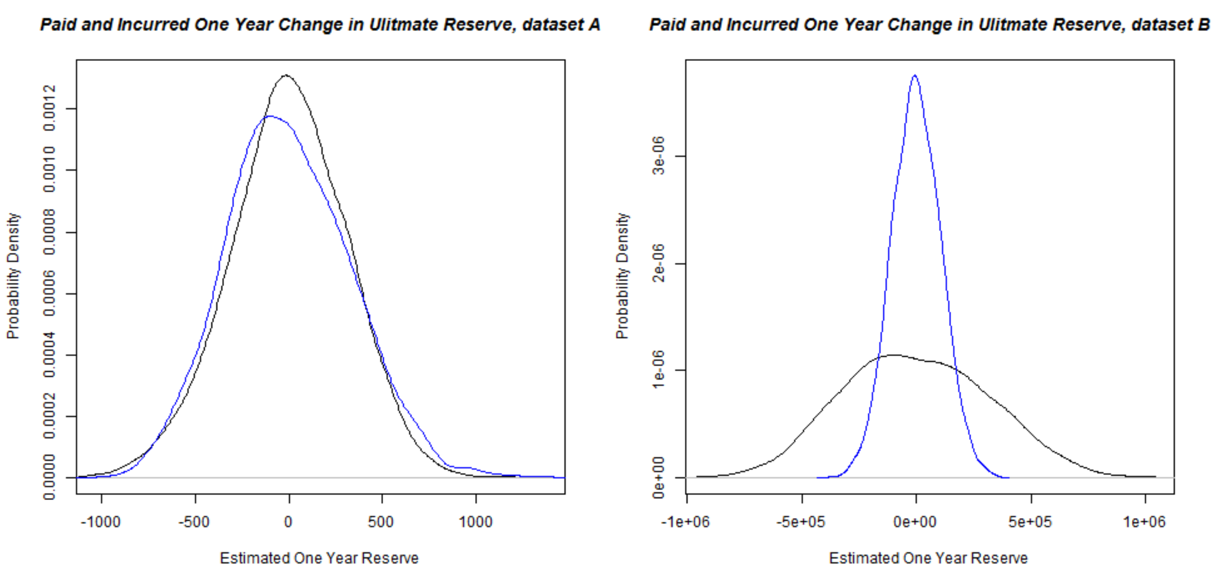

Figure 4.6: The PDF of OCDR, blue is from the incurred triangle and black is from the paid triangle.

Comparing Table 4.8 and 4.9 shows that for dataset A $\frac{\delta r_\infty}{\delta r_1}$ is equal to 1.37 and 1.18 for paid and incurred respectively, while for dataset 1 it is equal to 9.13 and 9.43 respectively.

	Dataset A	Dataset B	Dataset 1	Dataset 2	Dataset 3	Dataset 4
δr_1^{Paid}	0.109	0.0764	0.0677	0.327	0.148	0.0374
$\delta r_1^{Incurred}$	0.136	0.0270	0.0679	0.331	0.154	0.0404

Table 4.9: δr_1 for dataset A, B and 1 to 4.

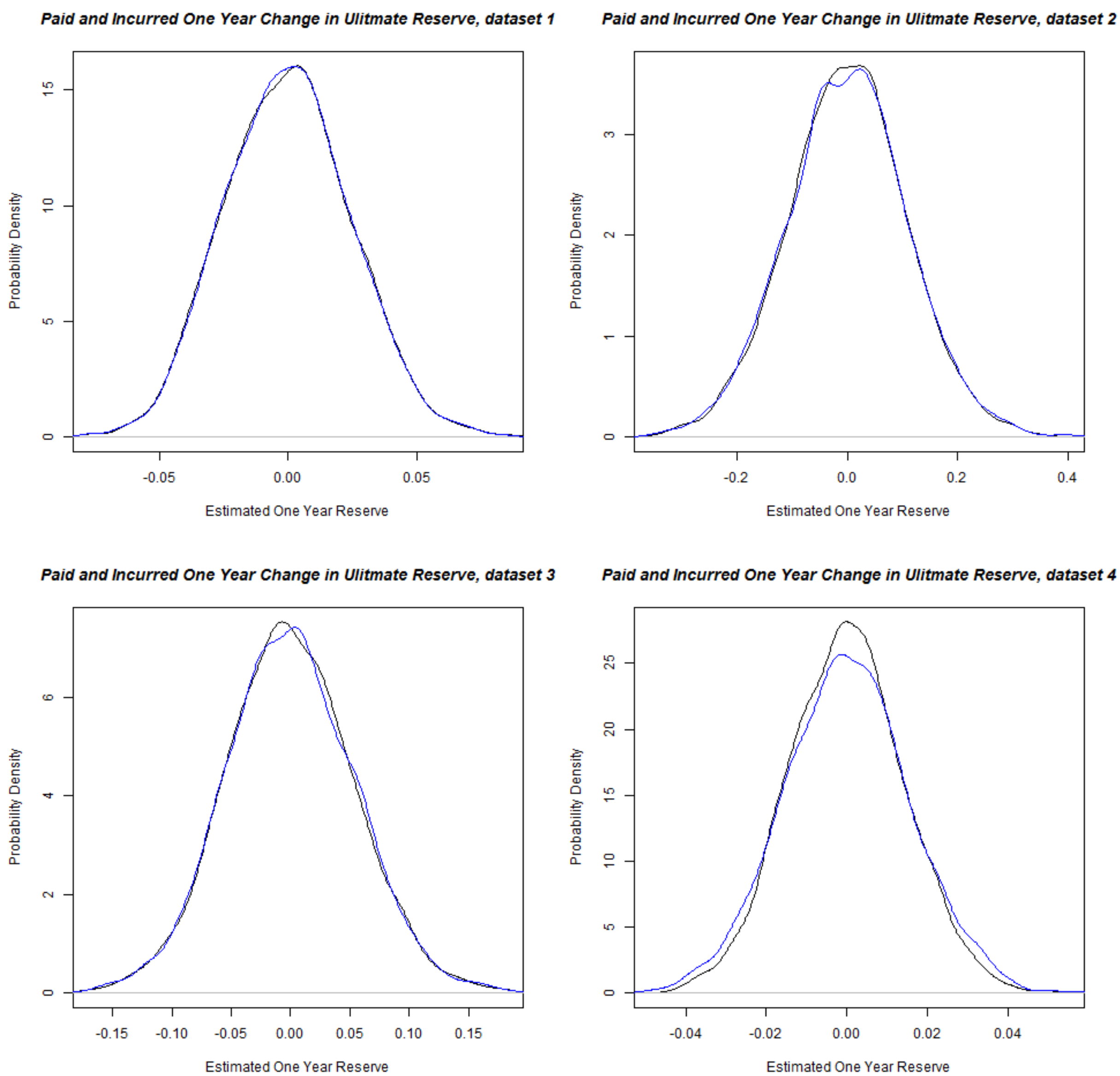

Figure 4.7: The PDF of OCDR, blue is from the incurred triangle and black is from the paid triangle.. The unit on the X-axis is $0.5R_\infty^{MCL,Paid} + 0.5R_\infty^{MCL,Incurred}$.

Chapter 5

Conclusions

In this thesis both MCL and BMCL, a bootstrapped MCL model were considered. With the results presented in this thesis the following conclusions can be drawn about MCL:

1. The predicted paid and incurred claims seems to converge for almost all datasets. This is true for all but one of the 38 insurance portfolios examined.
2. When looking at an arbitrary insurance portfolio, MCL seems to provide superior predictions compared to CLM with paid data in a small majority of the cases.
3. When only examning more recent accident years (after 1999) MCL seems to increase its relative predictive performance.
4. The relative performance of MCL compared to CLM seems to increase on average when only considering the datasets which fits the criteria of $\lambda^I \geq 0$, $\lambda^P \geq 0$ and no Q-value above 1.02
5. MCL does not seem to give any improvement on average over CLM with incurred data in predicting future claims. (Table 4.4 indicates that there could be a small improvement with datasets which fits the criteria of positive slopes and no Q-value above 1.02, however with n=14 it is hard to draw any conclusions.)

It was also shown that for some datasets MCL severely overcompensates when trying to converge the paid and incurred claims, leading to extreme divergence of the Q-values. These overcompensations can be alleviated by adding an upper limit to the Sigma-ratio.

The fact that the predicted paid and incurred claims converge means that when using MCL the choice of looking at future paid or incurred claims does not matter as much as when using CLM.

Conclusions two to five means that if one uses CLM with incurred data to calculate reserves there does not seem to be any great incentive to start using MCL instead. However if one is using CLM with paid data and if MCL have no negative λ for that dataset, then it could be advantageous to use MCL instead.

The fact that MCLs relative performance over CLM increases as the data is sorted, indicates that MCL is more sensitive to incomplete or incorrect data. This is expected as the MCL is a more complex method and as such there are many more ways an error in input data can propegate through the method.

With the results presented in this thesis the following conclusions can be drawn about BMCL:

1. An upper limit on the Sigma-ratio helps to alleviate overcompensation.
2. BMCL does have problems in that the mean of OCDR and the mean of estimated ultimate reserves do not converge to the wanted values (zero and the MCL ultimate reserve respectively).
3. Missing data, Q-value above 1.02, high Sigma-ratios and negative λ, all seems to increase the risk of a large difference between the wanted and the actual estimated means.
4. BMCL can be used to estimate both OCDR and the ultimate reserve PDF after the estimates mean values are set to the wanted values.
5. The PDF for paid and incurred reserves and risks estimated by BMCL seems to be very similar for most datasets, but not for all.
6. The one year risk estimated by BMCL is lower than the ultimate reserve risk estimated by BMCL, but how much lower is highly dependent on the dataset.

The fact that BMCL estimates do not converge to the right value is a potential weakness of the BMCL model and something that should be looked into. The fact that all the datasets with a 20 % or higher difference between the wanted and the actual estimated means had missing data, Q-values over 1.02, negative λ and/or high Sigma-ratio seems to indicate that BMCL is method that is vulnerable to datasets with these characteristics. Therefore perhaps one should not use BMCL to estimate the PDF for the ultimate reserve and OCDR for these datasets. However not all datasets with missing data, Q-values over 1.02, negative λ and/or high Sigma-ratio did have a large difference in the wanted and the actual estimated means.

The simulation error was not insignificant even at 10,000 simulations for dataset A and B, which indicates that if BMCL should be used in production it would be advantageous to use a higher number of simulations.

Bibliography

[1] Efron, B. (1979), *Bootstrap Methods: Another Look at the Jackknife*, The Annals of Statistics, Volume 7, Number 1, pages 1-26, January 1979

[2] Efron, B. and Tibshirani, R.J. (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall, 1993.

[3] England, P. D. (2002). *Addendum to "Analytic and bootstrap estimates of prediction errors in claims reserving"*, Insurance: Mathematics and Economics, Volume 31, Issue 3, December 2002, Pages 461-466

[4] Lauzeningks, J. and Ohlsson, E. (2008), *The one-year non-life insurance risk*, Ohlsson ASTIN Colloquium, July 2008

[5] Liu, H. and Verrall, R. (2010), *Bootstrap Estimation of the Predictive Distributions of Reserves Using Paid and Incurred Claims*, Variance, Volume 4, Issue 2, 2010, pages 121-135.

[6] Lysenko, N. , Merz, M. and Wüthrich, M. V. (2009). *Uncertainty in the claims development result in the chain ladder method*, Scandinavian Actuarial Journal, Volume 109, Number 1, 2009, Pages 63-84.

[7] Mack, T. (1993), *Distribution-free calculation of the standard error of chain ladder reserve estimates.*, ASTIN Bulletin, Volume 23:2, 1993, pages 213-225.

[8] Mack, T. and Quarg, G. (2004), *Munich Chain Ladder: A Reserving Method that Reduces the Gap between IBNR Projections Based on Paid Losses and IBNR Projections Based on Incurred Losses*, Variance, Volume 2, Number 2, 2008, pages 266-299.
First published in: Deutschen Gesellschaft für Versicherungs- und Finanzmathematik, Volume 26, number 4, 2004, pages 597-630.

[9] Merz, M. and Wüthrich, M.V. (2008), *Modelling the claims development result for solvency purposes*, CAS E-Forum, Fall 2008, pages 542-568, 2008.

[10] Merz, M. and Wüthrich, M. V. (2010), *Paid-Incurred chain claims reserving method*, Insurance: Mathematics and Economics, Volume 46, Issue 3, June 2010, pages 568-579.

[11] Miranda, M.D.M. , Nielsen, J.P. and Verrall, R. (2012), *Double chain ladder*, ASTIN Bulletin, Volume 42, Issue 1, 2012, pages 59-76

[12] Miranda, M.D.M. , Nielsen, J.P. , Verrall, R. and Wüthrich, M.V. (2015), *Double chain ladder, claims development inflation and zero-claims*, Scandinavian Actuarial Journal, Volume 2015, Issue 5, 2015, pages 383-405

[13] Renshaw, A. E. and Verrall, R. J. (1998) *A stochastic model underlying the chain-ladder technique.*, B.A.J., Volume 4, Issue 04, 1998, pages 903-923.

[14] Taylor, G. C. (1977), *Separation of Inflation and other Effects from the Distribution of Non-Life Insurance Claim Delays*, ASTIN Bulletin, Volume 9, Issue 1-2, January 1977, pages 219-230

[15] Verbeek, H. G. (1972), *An approach to the analysis of claims experience in motor liability excess of loss reinsurance.*, ASTIN Bulletin, Volume 6, Issue 3, 1977, pages 195-202