Stochastic claims reserving in non-life insurance

Stochastic claims reserving in non-life insurance

Bootstrap and smoothing models

Susanna Björkwall

Stockholm University

ISBN 978-91-7447-255-4

Printed in Sweden by US-AB, Stockholm 2011

Distributor: Department of Mathematics, Stockholm University

Abstract

In practice there is a long tradition of actuaries calculating reserve estimates according to deterministic methods without explicit reference to a stochastic model. For instance, the chain-ladder was originally a deterministic reserving method. Moreover, the actuaries often make ad hoc adjustments of the methods, for example, smoothing of the chain-ladder development factors, in order to fit the data set under analysis.

However, stochastic models are needed in order to assess the variability of the claims reserve. The standard statistical approach would be to first specify a model, then find an estimate of the outstanding claims under that model, typically by maximum likelihood, and finally the model could be used to find the precision of the estimate. As a compromise between this approach and the actuary's way of working without reference to a model the object of the research area has often been to first construct a model and a method that produces the actuary's estimate and then use this model in order to assess the uncertainty of the estimate. A drawback of this approach is that the suggested models have been constructed to give a measure of the precision of the reserve estimate without the possibility of changing the estimate itself.

The starting point of this thesis is the inconsistency between the deterministic approaches used in practice and the stochastic ones suggested in the literature. On one hand, the purpose of Paper I is to develop a bootstrap technique which easily enables the actuary to use other development factor methods than the pure chain-ladder relying on as few model assumptions as possible. This bootstrap technique is then extended and applied to the separation method in Paper II. On the other hand, the purpose of Paper III is to create a stochastic framework which imitates the ad hoc deterministic smoothing of chain-ladder development factors which is frequently used in practice.

"*...friends who suggested names more colorful than Bootstrap, including Swiss Army Knife, Meat Axe, Swan-Dive, Jack-Rabbit, and my personal favorite, the Shotgun, which, to paraphrase Tukey, "can blow the head off any problem if the statistician can stand the resulting mess"*."

Bradley Efron, 1979.
Bootstrap Methods: Another Look at the Jackknife.
The Annals of Statistics, vol. 7.

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.[1]

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

II Björkwall, S., Hössjer, O., Ohlsson, E. (2010) Bootstrapping the separation method in claims reserving. *ASTIN Bulletin*, 40(2): 845-869

III Björkwall, S., Hössjer, O., Ohlsson, E., Verrall, R. (2011) A generalized linear model with smoothing effects for claims reserving. *Insurance: Mathematics and Economics*, 49(1): 27-37.

S. Björkwall has contributed with the simulations, the analysis of the results and most of the writing, while the methodology was developed jointly.

[1] A related paper which is not included in the thesis:
Verrall, R., Hössjer, O., Björkwall, S. (2010) Modelling claims run-off with reversible jump Markov Chain Monte Carlo methods. Manuscript. Submitted.

Preface

When I was eight or nine years old I made a friend play insurance company with me. I was the actuary and she had to be the CEO. As an actuary I was doing my home work in math in one room and as the CEO she had to sit at a desk in another room reading and considering a file containing her parents bills. I guess that I do not even have to mention that she hated that game. I, on the other hand, really enjoyed it and I was very disappointed that I only could play it once.

It is quite strange - or perhaps it can be considered as totally expected - that I many years later actually ended up as an actuary and, moreover, that I now have written a doctoral thesis with actuarial applications. In any case, I am for all time grateful to those who gave me this opportunity. Therefore I would like to thank my supervisors Ola Hössjer and Esbjörn Ohlsson, who have made my wish come true by teaching me how to make scientific research of practical issues relating to my job. I am so glad that I finally found someone to share my interest with and I have had so much fun!

I also wish to thank Richard Verrall for the valuable cooperation which so far has resulted in two papers. I am deeply thankful for all the good advices, the inspiration and the opportunity to work with one of the greatest experts in my research area.

Finally, I would like to thank everyone else (family, friends, colleagues and idols) who has either helped me with practical things, supported me to reach my goal or showed interest in my research. This has motivated me to work even harder!

Susanna Björkwall
Stockholm, November, 2010

Contents

Part I: Introduction

1 Introduction 1
2 Claims reserving 3
2.1 Data 3
2.2 The chain-ladder method and modifications 6
2.3 The separation method 7
2.4 Stochastic claims reserving 9
2.5 Claims reserve uncertainty 10
3 Overview of papers 13
3.1 Paper I 13
3.2 Paper II 14
3.3 Paper III 15
4 Methodological discussion and concluding comments 17
4.1 Björkwall versus England 17
4.2 The standardized versus the unstandardized bootstrap 19
4.3 Prediction versus estimation 21
4.4 Research versus practical applications 21
5 Reserve risk in a business model 23
5.1 Solvency II 23
5.2 The one-year reserve risk 23
6 Some corrections 27
7 References 29

Part II: Papers

Part I:
Introduction

1. Introduction

Every non-life insurance company has to set up a fund for the future compensation of policy holders for claim events that have already occurred. This amount is usually referred to as the provision for outstanding claims or simply – the claims reserve. It is important that the claims reserve is carefully calculated; if it is underestimated the insurance company will not be able to fulfill its undertakings and if it is overestimated the insurance company unnecessarily holds the excess capital instead of using it for other purposes, e.g. for investments with higher risk and, hence, potentially higher return. Moreover, since the claims reserve usually constitutes a large share of the firm's total holdings even small miscalculations can imply considerable amounts of money.

On the basis of historical data the actuary can obtain estimates – or rather predictions – of the expected outstanding claims. However, due e.g. to poor data quality, or sometimes even lack of data, unexpectedly large claim payments, changes in inflation regime or in the discount rate and even legal and political factors, the uncertainty of the actuary's best estimate can be quite high. Obviously, there is a risk that the claims reserve will not suffice to pay all claims in the end or, in the one year perspective, that we get a negative run-off result. In order to monitor and manage this risk it is important that the actuary's best estimate is complemented by some measure of variability which can be followed up by the insurance company.

The literature provides a variety of methods for the actuary to choose amongst for reserving purposes, see e.g. the Claims Reserving Manual by the Faculty and Institute of Actuaries (1997). The reserving methods used in practice are frequently deterministic. For instance, the claims reserve is often obtained according to case estimation of individual claims by claims handlers. A popular statistical method is the chain-ladder method, see Taylor (2000), which originally was deterministic. Many ad hoc adjustments are applied as well, e.g. the projection of payments into the future can sometimes be done by extrapolating by eye. Hence, there is a long tradition of actuaries calculating reserve estimates without explicit reference to a stochastic model.

However, stochastic models are needed in order to assess the variability of the claims reserve. The standard statistical approach would be to first specify a model, then find an estimate of the outstanding claims under that model, typically by maximum likelihood, and finally the model could be used to find the precision of the estimate. As a compromise between this approach and the actuary's way of working without reference to a model the object of the research area called stochastic claims reserving has mostly been to first construct a model and a method that produces the actuary's best estimate and then use this model in order to assess the uncertainty of the estimate. In particular the object of several papers has been to find a model under which the best estimate is the one given by the chain-ladder method, see e.g. Verrall (2000), Mack & Venter (2000) and Verrall & England (2000).

Once the model has been chosen the variability of the claims reserve can be obtained either analytically or by simulation. For instance, the mean squared error of prediction (MSEP) for the chain-ladder method was first calculated analytically by Mack (1993). The reserve estimators are often complex functions of the observations and, hence, it might be difficult to derive analytical expressions. Therefore bootstrapping became a popular method when it was introduced for the chain-ladder by England & Verrall (1999) and England (2002). However, since the existing bootstrap techniques adopt the statistical assumptions in the literature, they have been constructed to give a measure of the precision of the actuary's best estimate without the possibility of changing the estimate itself.

The starting point of this thesis is the inconsistency between the deterministic approaches used in practice and the stochastic ones suggested in the literature. On one hand, the purpose of Paper I is to develop a bootstrap technique which easily enables the actuary to use other development factor methods than the chain-ladder relying on as few model assumptions as possible. This bootstrap technique is then extended and applied to the separation method, see Taylor (1977), in Paper II. On the other hand, the purpose of Paper III is to create a stochastic framework which imitates the ad hoc deterministic smoothing of chain-ladder development factors which is frequently used in practice in order to obtain more reliable and robust reserve estimates.

The thesis summary is set out as follows. Chapter 2 provides an introduction to a practical reserving exercise as well as to the methods discussed and developed in Papers I-III. An overview of Papers I-III is given in Chapter 3, while the chosen methodologies are commented and analyzed further in Chapter 4. Finally, Chapter 5 demonstrates how bootstrapping could be used in order to estimate the reserve risk in practice.

2. Claims reserving

2.1 Data

Large insurance companies often have quite extensive data bases with historical information on incurred claims. Such information can include the numbers of claims reported and settled, the origin year of the events, the paid amounts, the year of the payments and case estimates. The actuary can regularly analyze the data in order to predict the outstanding claims and, hence, the claims reserve.

The analysis is typically done in the following way. To begin with, the actuary separates the data into risk homogenous groups such as lines of business, e.g. Motor, Property and Liability. A finer segmentation can be applied if the groups or the subgroups contain a sufficient number of observations. The actuary might also choose to divide some group according to the severity of the claims. The large claims can then be reserved according to case estimates while the subgroup consisting of smaller, but frequently occurring, claims can be reserved by some statistical method.

When the risk classification is established the actuary usually aggregates the data within the groups into development triangles. We now consider such an incremental triangle of paid claims $\{C_{ij}; i, j \in \nabla\}$, where the business has been observed during t years, i.e. $\nabla = \{i = 0, \ldots, t; j = 0, \ldots, t-i\}$[1]. The suffixes i and j of the paid claims refer to the origin year and the payment year, respectively, see Table 2.1. In addition, the suffix $k = i + j$ is used for the calendar years, i.e. the diagonals of ∇.

If we assume that the claims are settled within the t observed years the purpose of a claims reserving exercise is to predict the sum of the delayed claim amounts in the lower, unobserved future triangle $\{C_{ij}; i, j \in \Delta\}$, where $\Delta = \{i = 1, \ldots, t; j = t-i+1, \ldots, t\}$, see Table 2.2. We write $R = \sum_{\Delta} C_{ij}$ for this sum, which is the outstanding claims for which the insurance company must hold a reserve. The outstanding claims per origin year are specified by $R_i = \sum_{j \in \Delta_i} C_{ij}$, where Δ_i denotes the row corresponding to origin year i in Δ.

Estimators of the outstanding claims per origin year and the grand total are obtained by $\hat{R}_i = \sum_{j \in \Delta_i} \hat{C}_{ij}$ and $\hat{R} = \sum_{\Delta} \hat{C}_{ij}$, respectively, where $\hat{C}_{ij}$ is a prediction of C_{ij}. With an underlying stochastic reserving model, $\hat{C}_{ij}$ is a function of the estimated parameters of that model, typically chosen to make it an

[1]Note that Paper II uses this notation, while Paper I and Paper III use $\nabla = \{i = 1, \ldots, t; j = 1, \ldots, t-i\}$

Origin year	Development year 0	1	2	$\cdots$	$t-1$	t
0	C_{00}	C_{01}	C_{02}	$\cdots$	$C_{0,t-1}$	$C_{0,t}$
1	C_{10}	C_{11}	C_{12}	$\cdots$	$C_{1,t-1}$	
2	C_{20}	C_{21}	C_{22}	$\cdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$t-1$	$C_{t-1,0}$	$C_{t-1,1}$				
t	$C_{t,0}$					

Table 2.1: The triangle ∇ of observed incremental payments.

(asymptotically) unbiased predictor of C_{ij}. However, as indicated in Chapter 1, algorithms not based on a probability model are often used in practice in order to obtain a reliable value of $\hat{C}_{ij}$.

Some reserving methods consider additional information. Hence, we also assume that the actuary can sum up a triangle of incremental observations of the numbers of claims $\{N_{ij}; i, j \in \nabla\}$ corresponding to the same portfolio as in Table 2.1, i.e. the observations in Table 2.3. The ultimate number of claims relating to the period of origin year i is then

$$N_i = \sum_{j \in \nabla_i} N_{ij} + \sum_{j \in \Delta_i} N_{ij}\,, \tag{2.1}$$

where ∇_i denotes the row corresponding to origin year i in the upper triangle ∇.

Origin year	Development year 0	1	2	$\cdots$	$t-1$	t
0						
1						$C_{1,t}$
2					$C_{2,t-1}$	$C_{2,t}$
$\vdots$					$\vdots$	$\vdots$
$t-1$			$C_{t-1,2}$	$\cdots$	$C_{t-1,t-1}$	$C_{t-1,t}$
t		$C_{t,1}$	$C_{t,2}$	$\cdots$	$C_{t,t-1}$	$C_{t,t}$

Table 2.2: The triangle Δ of unobserved future claim costs.

When the paid amounts are presented as in Table 2.1 the payment pattern emerges along the rows, while the columns indicate the size of the business

Origin year	Development year 0	1	2	$\cdots$	$t-1$	t
0	N_{00}	N_{01}	N_{02}	$\cdots$	$N_{0,t-1}$	$N_{0,t}$
1	N_{10}	N_{11}	N_{12}	$\cdots$	$N_{1,t-1}$	
2	N_{20}	N_{21}	N_{22}	$\cdots$		
$\vdots$	$\vdots$	$\vdots$	$\vdots$			
$t-1$	$N_{t-1,0}$	$N_{t-1,1}$				
t	$N_{t,0}$					

Table 2.3: The triangle ∇ of observed incremental numbers of reported claims.

over the origin years. Moreover, the diagonals show calendar year effects. Hence, regularities as well as irregularities become apparent to the actuary. For instance, occurrence of growth or decrease of the business, claims inflation or rare large claims can usually be detected in the development triangle and the actuary can then decide how to deal with these issues. If the business is growing or decreasing the actuary can disregard the earliest origin years which might have another payment pattern than the later ones. In case of inflation the payments can be adjusted to current value by some relevant index or a reserving method which considers inflation can be chosen. Claims originating from large events and catastrophes can be excluded from the triangle and treated separately.

Note that if observations are missing for some years the data in Table 2.1 will have another shape. Henceforth we assume that the data has the shape of a complete triangle. However, despite a complete triangle the information can still be insufficient if the business has not been observed during a sufficient time period. This is usually a problem for long-tailed lines of business, such as Motor TPL, where it can take several decades to settle the claims. We then have no origin year with finalized claims in Table 2.1. When needed, the reserving methods can be extended so that the unknown claims extend beyond t in a tail of length u, i.e. over the development years $t+1,\dots,t+u$.

It is worth bearing in mind that sometimes the data quality may be increased and the reserving process may be refined, but only at a cost. In practice the amount of time and the cost of improving the processes have to be related to the benefits, but even if faster and cheaper approximations are chosen it is still important that the actuary is aware of e.g. imperfections in the data and how they affect the results.

2.2 The chain-ladder method and modifications

The chain-ladder method is probably the most popular reserving technique in practice. According to Taylor (2000) its lineage can be traced to the mid-60's and the name should refer to the chaining of a sequence of age-to-age development factors into a ladder of factors by which one can climb from the observations to date to the predicted ultimate claim cost. The chain-ladder was originally deterministic, but in order to assess the variability of the estimate it has been developed into a stochastic method. Taylor (2000) presents different derivations of the chain-ladder procedure; one of them is deterministic while another one is based on the assumption that the incremental observations are Poisson distributed. Verrall (2000) provides several models which under maximum likelihood estimation reproduce the chain-ladder reserve estimate.

The chain-ladder method operates on cumulative observations

$$D_{ij} = \sum_{\ell=0}^{j} C_{i\ell} \tag{2.2}$$

rather than incremental observations C_{ij}.

Let $\mu_{ij} = E(D_{ij})$. Development factors

$$f_j = \frac{\sum_{i=0}^{t-j-1} \mu_{i,j+1}}{\sum_{i=0}^{t-j-1} \mu_{ij}}, \tag{2.3}$$

where $j = 0, \ldots, t-1$, are estimated for a fully non-parametric model by

$$\hat{f}_j = \frac{\sum_{i=0}^{t-j-1} D_{i,j+1}}{\sum_{i=0}^{t-j-1} D_{ij}}. \tag{2.4}$$

Projections can then be obtained by

$$\hat{\mu}_{ij} = D_{i,t-i}\, \hat{f}_{t-i}\, \hat{f}_{t-i+1} \ldots \hat{f}_{j-1} \tag{2.5}$$

and

$$\hat{C}_{i,j} = \hat{\mu}_{i,j} - \hat{\mu}_{i,j-1} \tag{2.6}$$

for Δ.

Note that the chain-ladder can be used in order to estimate any other quantity of interest too, e.g. the ultimate number of claims N_i in (2.1). In that case C_{ij} is substituted by the number of claims N_{ij} in the notation above.

The actuary might want to make some ad hoc adjustments of the chain-ladder method in order to deal with the trends and occurrences of the influences discussed in Section 2.1. The reserving method is then usually referred to as an age-to-age development factor method and since it will be unique for the particular data set under analysis it is impossible to describe it in general terms. Here is an example of an adjusted procedure that might fit our scheme when ∇C is available:

Firstly, the chain-ladder is used to produce development factors $\hat{f}_j$ according to (2.4), perhaps after excluding the oldest observations and/or sole outliers in ∇C.

By examining a graph of the sequence of $\hat{f}_j$'s the actuary might decide to smooth them for say $j \geq 4$. Exponential smoothing could be used for that purpose, i.e. the $\hat{f}_j$'s are replaced by estimates obtained from a linear regression of $\ln(\hat{f}_j - 1)$ on j. By extrapolation in the linear regression this also yields development factors for a tail $j = t, t+1 \ldots, t+u-1$. The original $\hat{f}_j$'s are kept for $j < 4$ and the smoothed ones used for all $j \geq 4$.

Let $\hat{f}_j^s$ denote the new sequence of development factors. Estimates $\hat{\mu}_{ij}$ for Δ are then computed as in the standard chain-ladder method yielding

$$\hat{\mu}_{ij} = D_{i,t-i}\, \hat{f}_{t-i}^s\, \hat{f}_{t-i+1}^s \cdots \hat{f}_{j-1}^s \tag{2.7}$$

and

$$\hat{C}_{i,j} = \hat{\mu}_{i,j} - \hat{\mu}_{i,j-1}\,. \tag{2.8}$$

The obtained claim values may be discounted by some interest rate curve or inflated by assumed claims inflation. The latter of course requires that the observations were recalculated to fixed prices in the first place.

Finally, an estimator $\hat{R} = h(\nabla C)$ can be obtained for some possibly quite complex function h which might be specified only by an algorithm as in the example above.

This approach might be considered as somewhat ad hoc from a statistical point of view, since it is difficult to adopt in the context of maximum likelihood estimation. Moreover, it is also difficult to systematize the procedure since e.g. the truncation point $j = 4$ of the unsmoothed development factors has to be decided by eye.

2.3 The separation method

In the Encyclopedia of Actuarial Science by Teugels & Sundt (2004) one can read that the separation method was developed by Taylor (1977) while he was employed at the Department of Trade, the supervisory authority in the UK. During the mid-70's the inflation was high and unstable and the Department of Trade had been experimenting with the inflation-adjusted version of the chain-ladder, see e.g. Taylor (2000). However, the specification of the future inflation caused problems, since it was extremely controversial for a supervisory tool. As an attempt to forecast the inflation mechanically Taylor (1977) constructed the separation method on the basis of a technique introduced in the reinsurance context by Verbeek (1972).

The separation method was, like the chain-ladder, originally formulated without detailed distributional assumptions. The major difference between the two methods is that the chain-ladder only makes implicit allowance for claims inflation since it projects the inflation present in the past data into the future, while the separation method incorporates it into the model underlying the reserving method. The former approach only works properly when the inflation rate is constant.

The original assumption underlying the separation method is

$$E\left(\frac{C_{ij}}{N_i}\right) = r_j \lambda_k , \tag{2.9}$$

where r_j is a parameter relating to the payment pattern for the development years, while λ_k is considered as an index that relates to the calendar year k during which the claims are paid. In this way the separation method separates the claim delay distribution from influences affecting the calendar years, e.g. claims inflation. Furthermore, it is assumed that the claims are fully paid by year t and we then have the constraint

$$\sum_{j=0}^{t} r_j = 1. \tag{2.10}$$

If N_i is estimated separately, e.g. by the chain-ladder if a triangle of claim counts is provided, it can be treated as known. Consequently, estimates $\hat{r}_j$ and $\hat{\lambda}_k$ can be obtained using the observed values

$$s_{ij} = \frac{C_{ij}}{\hat{N}_i} \tag{2.11}$$

and the marginal sum equations

$$s_{k0} + s_{k-1,1} + \ldots + s_{0k} = (\hat{r}_0 + \ldots + \hat{r}_k)\,\hat{\lambda}_k , \qquad k = 0, \ldots, t \tag{2.12}$$

for the diagonals of ∇ and

$$s_{0j} + s_{1j} + \ldots + s_{t-j,j} = (\hat{\lambda}_j + \ldots + \hat{\lambda}_t)\,\hat{r}_j , \qquad j = 0, \ldots, t \tag{2.13}$$

for the columns of ∇.

Taylor (1977) shows that the equations (2.12) - (2.13) have a unique solution under (2.10) which can be obtained recursively. This yields

$$\hat{\lambda}_k = \frac{\sum_{i=0}^{k} s_{i,k-i}}{1 - \sum_{j=k+1}^{t} \hat{r}_j} , \qquad k = 0, \ldots, t \tag{2.14}$$

and

$$\hat{r}_j = \frac{\sum_{i=0}^{t-j} s_{ij}}{\sum_{k=j}^{t} \hat{\lambda}_k} , \qquad j = 0, \ldots, t , \tag{2.15}$$

where $\sum_{j=k+1}^{t} \hat{r}_j$ is interpreted as zero when $k = t$.

Estimates $\hat{C}_{ij}$ for cells in ∇ can now be computed by

$$\hat{C}_{ij} = \hat{N}_i \hat{r}_j \hat{\lambda}_k , \tag{2.16}$$

but in order to obtain the estimates of Δ it remains to predict λ_k for $t+1 \leq k \leq 2t$. This can be done e.g. mechanically by extrapolation.

2.4 Stochastic claims reserving

The chain-ladder as well as the separation method can be considered as deterministic methods which are described as mechanical algorithms rather than full models. The advantage of this approach is that it is intuitive and transparent even for non-actuaries who are involved in the reserving process. However, the statistical quality of these reserve estimates cannot be assessed unless a model is found for which the algorithms are, for instance, obtained as functions of maximum likelihood estimators, nor is it possible to get a measure of the uncertainty of the reserve estimates.

Many stochastic models have been established within the research area of stochastic claims reserving, see, for example, Wütrich & Merz (2008) and England & Verrall (2002) for a summary. A popular stochastic model is the generalized linear model (GLM) introduced in a claims reserving context by Renshaw & Verrall (1998). A common assumption in the literature is to use an over-dispersed Poisson (ODP) distribution and a logarithmic link function for the incremental observations ∇C in Table 2.1. A consequence of this particular assumption is that the estimates of the expected claims obtained by maximum likelihood estimation of the parameters in the GLM equal the ones obtained by the chain-ladder method, if the column sums of the triangle are positive, see Renshaw & Verrall (1998). Thus, the estimated expected values can be obtained either by maximum likelihood estimation or by the chain-ladder and the estimated variances, which are obtained from the GLM assumption, could be used in order to derive or simulate a variability measure.

The following Tweedie GLM with a log-link will be frequently used throughout this thesis

$$\begin{aligned} E(C_{ij}) = m_{ij} \quad &\text{and} \quad Var(C_{ij}) = \phi_{ij}\, m_{ij}^{p} \\ \ln(m_{ij}) \quad &= \quad \eta_{ij}\,, \end{aligned} \tag{2.17}$$

where

$$\eta_{ij} = c + \alpha_i + \beta_j\,. \tag{2.18}$$

Here, the scale parameter is usually assumed to be constant, i.e. $\phi_{ij} = \phi$. Equation (2.18) can be extended to include a calender year parameter according to

$$\eta_{ij} = c + \alpha_i + \beta_j + \gamma_k, \qquad k = 0, \ldots, 2t\,, \tag{2.19}$$

however, the number of parameters is then usually too large compared to the small data set of aggregated individual paid claims. In any case, a constraint, e.g.

$$\alpha_0 = \beta_0 = \gamma_0 = 0\,, \tag{2.20}$$

is needed to estimate the remaining model parameters c, α_i, β_j and γ_k, typically under the assumption $p = 1$ or $p = 2$, corresponding to an ODP distribution or a gamma distribution, respectively. Note that it is only possible to estimate

γ_k for $k = 1, \ldots, t$, while a further assumption is needed regarding the future $k = t+1, \ldots, 2t$.

Many stochastic models which, in particular, considers smoothing of the run-off pattern can be found in the literature. A source of inspiration for Paper III is the Generalized Additive Model (GAM) framework presented in England & Verrall (2001). This framework has the flexibility to include several well-known reserving models as special cases as well as to incorporate smoothing and extrapolation in the model-fitting procedure. Using the framework implies that the actuary simply would have to choose one parameter corresponding to the amount of smoothing, the error distribution and how far to extrapolate, then the fitted model automatically provides statistics of interest, e.g. reserve estimates and measures of precision.

In addition to (2.17), England & Verrall (2001) use the GAM

$$\eta_{ij} = u_{ij} + \delta k + c + s_{\theta_i}(i) + s_{\theta_j}(j) + s_{\theta_j}(\ln(j)), \tag{2.21}$$

where the functions $s(i)$, $s(j)$ and $s(\ln(j))$ represent smoothers on i, j and $\ln(j)$, respectively, using the smoothing parameters θ_i and θ_j. Here $s_0(i) = \alpha_i$ corresponds to no smoothing, while $s_\infty(i)$ implies full smoothing, i.e. a linear function of i. The corresponding holds for j. The offsets u_{ij}, which are known terms in the particular context, and δk, corresponding to inflation, in (2.21) are optional.

2.5 Claims reserve uncertainty

An analytical derivation of the prediction error of $\hat{R}$ may be preferable from a theoretical perspective, however, this approach is often impracticable due to complex reserve estimators. Early attempts were essentially based on least squares regression applied to the logarithms of the incremental observations ∇C in Table 2.1, see e.g. Zehnwirth (1989) and Verrall (1991).

For the chain-ladder method, Mack (1993) derived an analytical expression of the MSEP within an autoregressive formulation of the claims development using a second-moment assumption. The procedure is exact, but it only holds for the pure chain-ladder. Hence, any adjustments of the reserve estimator, e.g. as described in Section 2.2, would require a change of the underlying assumptions. Moreover, an additional distributional assumption is needed in order to obtain a full predictive distribution.

Renshaw (1994) and England & Verrall (1999) derived a first order Taylor approximation of the corresponding MSEP within the GLM framework in (2.17) and (2.18). They found that the MSEP can be decomposed as

$$MSEP(R) \approx Var(R) + Var(\hat{R}). \tag{2.22}$$

Consequently,

$$MSEP(R) \approx \sum_{i,j\in\Delta} \phi m_{ij}^{p} + \sum_{i,j\in\Delta} m_{ij}^{2} \mathrm{Var}(\hat{\eta}_{ij}) + \sum_{\substack{(i_1,j_1),\\(i_2,j_2)\in\Delta\\ i_1 j_1 \neq i_2 j_2}} m_{i_1 j_1} m_{i_2 j_2} Cov(\hat{\eta}_{i_1 j_1}, \hat{\eta}_{i_2 j_2}) \tag{2.23}$$

could be derived, since the process error component is

$$Var(R) = \sum_{\Delta} \phi\, m_{ij}^{p} \tag{2.24}$$

under the assumed model. In order to calculate the matrix of the covariance terms, England & Verrall (1999) extracted the design matrix and variance-covariance matrix of the parameter estimates from the statistical software.

England & Verrall's derivation of (2.23) using the GLM assumption clearly demonstrates the strength of a stochastic model. However, even though (2.23) is particularly true for the model in Paper III it may be difficult to derive Taylor series expansions for other extensions of the chain-ladder method, and further no predictive distribution of the reserve is obtained.

England & Verrall (1999) extended equation (2.23) by suggesting the use of bootstrapping to obtain the estimation error component $Var(\hat{R})$ in (2.22). When $p = 1$ they replace (2.22) by

$$\widehat{MSEP}(\hat{R}) \approx \hat{\phi}\hat{R} + \widehat{Var}(\hat{R}^{*}), \tag{2.25}$$

where $\widehat{Var}(\hat{R}^{*})$ is the variance of simulated values of $\hat{R}^{*}$ obtained by a bootstrap procedure.

The suggested bootstrap procedure involves resampling of the adjusted Pearson residuals[2]

$$r_{ij} = \sqrt{\frac{n}{n-q}} \frac{C_{ij} - \hat{m}_{ij}}{\sqrt{\hat{m}_{ij}}}, \tag{2.26}$$

where n is the number of observations in ∇C and q is the number of estimated parameters. Hence, B bootstrap samples ∇r^{*} are generated, which then are converted to pseudo-triangles ∇C^{*} by computing

$$C_{ij}^{*} = \hat{m}_{ij} + r_{ij}^{*}\sqrt{\hat{m}_{ij}} \tag{2.27}$$

for $i, j \in \nabla$. The future values $\Delta\hat{m}^{*}$ of the B pseudo-triangles are forecasted by the chain-ladder and finally $\hat{R}^{*} = \sum_{\Delta} \hat{m}_{ij}^{*}$ is obtained.

In order to simulate a full predictive distribution England (2002) extended the method in England & Verrall (1999) by replacing the analytic calculation of the process error by another simulation conditional on the bootstrap simulation. The process error is added to the B triangles $\Delta\hat{m}^{*}$ by sampling random

[2]The adjustment term is needed in order to account for the number of parameters used in fitting the model. This may be accounted for in the variance and covariance terms of (2.23).

observations from distributions with mean $\hat{m}^*_{ij}$ and variance $\hat{\phi}\hat{m}^*_{ij}$ to obtain the future claims $\Delta m^\dagger$. The predictive distribution of the outstanding claims is then obtained by plotting the B values of $R^\dagger = \sum_\Delta m^\dagger_{ij}$.

The bootstrap approach suggested by England (2002) is not consistent with the plug-in-principle, see Efron & Tibshirani (1993), since the second simulation stage proceeds conditionally on the first bootstrap sampling. Consequently, this approach does not allow for adjustments of the reserve estimator. In contrast to England & Verrall (1999) and England (2002), Pinheiro *et al.* (2003) adopts the model in (2.17) and (2.18) together with the plug-in-principle. Hence, the relation between the true outstanding claims R and its estimator $\hat{R}$ in the *real world* is, by the plug-in-principle, substituted in the *bootstrap world* by their bootstrap counterparts. This implies that the process error is simulated separately from the estimation error; the former is included in R^{**}, i.e. the true outstanding claims in the bootstrap world, while the latter is included in $\hat{R}^*$, i.e. the estimated outstanding claims in the bootstrap world. Here, $\hat{R}^*$ is obtained as in England (2002) and for $R^{**} = \sum_\Delta C^{**}_{ij}$, the resampling is done once more to get B triangles of Δr^{**} and then solving

$$C^{**}_{ij} = \hat{m}_{ij} + r^{**}_{ij}\sqrt{\hat{m}_{ij}} \quad \text{for} \quad i,j \in \Delta \tag{2.28}$$

to get ΔC^{**}. Pinheiro *et al.* (2003) suggest the use of a standardized prediction error, however, the unstandardized $\mathrm{pe}^{**} = R^{**} - \hat{R}^*$ yields $\tilde{R}^{**} = \hat{R} + \mathrm{pe}^{**}$, which is comparable to England's $R^\dagger$.

Pinheiro *et al.* (2003) generalize the bootstrap procedure to allow for $p = 2$ in (2.17). However, note that the underlying GLM and the suggested bootstrap approaches in England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) are chosen particularly in order to obtain procedures for the chain-ladder method.

3. Overview of papers

3.1 Paper I

When England & Verrall (1999) and England (2002) introduced bootstrapping in claims reserving it soon became a popular method in practice as well as in the literature. However, even though bootstrapping has been hailed as a flexible tool to find the precision of the complex reserve estimators it has continued to be the opposite in the literature. Instead of finding general techniques where the actuary can change and adjust the reserving method, the object of the research area has often been to find techniques for, in particular, the chain-ladder. In practice this could be quite frustrating since the actuary then has to measure the uncertainty of her estimate by a bootstrap procedure fitted for the pure chain-ladder even though she actually has used some other reserving method to calculate the claims reserve.

Therefore, the purpose of this paper is to relax the model assumption in England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) in order to obtain a bootstrap approach which could be used for other development factor methods than the chain-ladder. Since Pinheiro *et al.* (2003) adopt the plug-in-principle, which theoretically enables adjustments in the reserve estimator, we will focus on extending Pinheiro's method.

We consider the log-additive assumption in (2.18) as unnecessary strong for reserving purposes. Besides of that we continue to follow England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) assuming independent claims C_{ij} and a variance function in terms of the means, i.e.

$$E(C_{ij}) = m_{ij} \quad \text{and} \quad \mathrm{Var}(C_{ij}) = \phi\, m_{ij}^{p} \tag{3.1}$$

for some $p > 0$. We let the actuary's age-to-age development factor method implicitly specify the structure of all m_{ij} and produce estimates of $\hat{m}_{ij}$. Then, if the non-parametric bootstrap approach of Pinheiro *et al.* (2003) is used, it only remains to specify the variance function. Instead of assuming $p = 1$ in order to reproduce the chain-ladder estimates, we suggest that p is estimated from data and we provide a simple and straightforward way of doing it. Moreover, since the standardized prediction errors in Pinheiro *et al.* (2003) sometimes are undefined in the bootstrap world we also investigate a bootstrap procedure which is based on unstandardized prediction errors.

As a complement to Pinheiro's non-parametric bootstrap approach we define a parametric version that requires more distributional assumptions. Hence, instead of resampling the residuals we directly sample pseudo-observations from a full distribution $F = F(m_{ij}, \phi\, m_{ij}^{p})$ consistently with (3.1). This approach specfically suits small data sets, since few observations generate an inadequate empirical distribution of the residuals.

The numerical study shows that the suggested bootstrap approaches are viable alternatives to the existing ones. It also demonstrates that the estimation error component is much larger than the process error component under the chosen reserving method and variance assumption. This is a characteristic of claims reserving since the actuary attemps to predict the future based on tiny sets of historical observations. It is by no means clear that estimation error should be relatively smaller for large triangles, in spite of the fact that there is more data available for a large triangle, since parameters corresponding to late origin and development years are hard to estimate for large as well as small development triangles.

3.2 Paper II

Paper II is a continuation of Paper I and the purpose is to make use of the flexibility of the suggested bootstrap approach. In particular, the purpose is to measure the uncertainty in the reserve estimate obtained by the separation method, which also considers calendar year effects.

In order to implement a bootstrap procedure for the deterministic separation method it has to be given a stochastic formulation and, moreover, the bootstrap approach has to be extended to handle ∇N as well as ∇C. To this end, we introduce a parametric framework where claim counts are Poisson distributed and claim amounts are gamma distributed *conditionally* on the ultimate claim counts. This enables joint resampling of claim counts and claim amounts.

Hence, we let $n_{ij} = E(N_{ij})$ and assume

$$N_{ij} \in Po(n_{ij}) \tag{3.2}$$

and

$$C_{ij}|N_i \in \Gamma\left(\frac{N_i}{\phi}, r_j\, \lambda_k\, \phi\right). \tag{3.3}$$

We then get a model for the claim amounts where

$$E(C_{ij}|N_i) = N_i\, r_j\, \lambda_k\,, \tag{3.4}$$

which is consistent with the separation method assumption (2.9) when N_i is estimated separately. Moreover, we have

$$Var(C_{ij}|N_i) = \phi\, N_i\, (r_j\, \lambda_k)^2\,. \tag{3.5}$$

The separation method requires that the inflation rates λ_k are predicted for $t+1 \leq k \leq 2t$ and in this paper $\lambda_k = (1+K)^k\, \lambda_0$ is used. The future inflation rates can of course be modeled by more refined approaches, but this is beyond the scope of Paper II and, hence, two simple models are considered. The first one is to use the mean rate observed so far, i.e. $\hat{K}_{mean}$, and the second one is to estimate K by loglinear regression, i.e. $\hat{K}_{reg}$.

The predictive distribution can be assessed using an extension of the parametric bootstrap procedure in Paper I. Hence, (3.2) and (3.3) are used in order to generate pseudo-observations ∇N^* and ∇C^*, while the plug-in-principle is used to produce estimates of r_j and λ_k in the bootstrap world analogously as in the real world. Moreover, the future inflation rates are predicted for each pseduo-triangle in the bootstrap world. The process error is generated in the second sampling stage by simulating ΔN^{**} as well as ΔC^{**}.

As expected, the numerical study shows that the claims inflation contributes to the uncertainty in the reserve estimate. Hence, it is important to consider its impact in the context of risk management. Nevertheless, inflation tend to be disregarded in practice, since it might be considered as an over-parametrization of the model relative to the small triangles.

3.3 Paper III

In contrast to the previous papers, the purpose of Paper III is to create a stochastic framework for a reserving exercise which is performed somewhat arbitrarily in practice.

It is necessary for stochastic reserving models to enable intuitive adjustments, for example, smoothing of the shape of the development pattern as discussed in Section 2.2. In order to implement such a smoothing model we suggest the use of a reparameterized version of (2.19). This model is already popular in a claims reserving context and a reparameterized version enables smoothing of origin, development and calendar year parameters in a similar way as is often done in practice. In this way the GLM structure is kept and it can be used to obtain reserve estimates and to systemize the model selection procedure that arises in the smoothing process.

Paper III provides a model which considers log-linear smoothing of all the parameters α_i, β_j and γ_k in (2.19), but here we only summarize the method for the shape of the development pattern. Hence, we suggest the use of the following model[1]

$$\begin{aligned} \beta_j &= b_{j-1}; \quad 1 \leq j \leq r \\ \beta_j &= b_{r-1} + b_r\,(j-r); \quad r+1 \leq j \leq t\,, \end{aligned} \tag{3.6}$$

where $b = \begin{pmatrix} b_0 & \dots & b_r \end{pmatrix}$ is a new set of parameters and $0 \leq r \leq t-1$.

Model (3.6) implies that the original parameters β_j are kept up to a certain truncation point r, thereafter we instead use a fitted linear curve in a similar

[1]Note that the indexation in Paper III differs from the one introduced in Section 2.1.

way as was described in Section 2.2. However, now the question of how r should be chosen arises.

Let $\hat{\theta}_r$ denote the estimated parameter vector for a model with a fixed r. We can choose the model

$$\hat{r} = \operatorname{argmin}_{r \in I} \mathrm{Crit}(\hat{\theta}_r), \tag{3.7}$$

that minimizes a model selection criterion $\mathrm{Crit}(\hat{\theta}_r)$ among a pre-chosen set I of candidate models. We then take $\hat{\theta}_{\hat{r}}$ as the final parameter estimate on which to base reserves.

Three different selection critera are investigated. Akaike's Information Criterion (AIC)

$$\mathrm{Crit} = \mathrm{AIC}(\hat{\theta}_r) = 2w - 2l(\hat{\theta}_r) \tag{3.8}$$

and the Bayesian Information Criterion (BIC)

$$\mathrm{Crit} = \mathrm{BIC}(\hat{\theta}_r) = \ln(n)w - 2l(\hat{\theta}_r) \tag{3.9}$$

are used when inference is based on likelihood functions. Here $w = 1 + r$ is the number of parameters and $l(\hat{\theta}_r)$ is the maximized log-likelihood function with respect to model r.

Bootstrapping provides our third criterion through

$$\mathrm{Crit} = \widehat{MSE}P(\hat{\theta}_r) = E\left((R^{**} - \hat{R}^*)^2\right), \tag{3.10}$$

where the resampled data are created by a parametric bootstrap from model r.

The numerical study shows that the distributional assumption of the model had a larger impact on the results than the smoothing effect. Hence, it seems important to first find an appropriate model, which then possibly could be adjusted by smoothing of the model parameters.

4. Methodological discussion and concluding comments

4.1 Björkwall versus England

The purpose of Paper I was to develop a bootstrap procedure which could be used in subsequent papers. The method introduced in England (2002) was excepted at an early stage, partly since it does not support the use of the plug-in-principle and partly since it was argued that this approach does not provide the right predictive distribution, see Appendix B in Paper I. Consequently, it was outside the scope of the paper to numerically compare the suggested procedures to England's approach.

However, among actuaries England's bootstrap method is well-known and frequently used to obtain an estimate of the reserve uncertainty. A crucial reason for its popularity is the simplicity and straightforwardness of the implementation. Even though it, in theory, could be argued that the approach is somewhat entangled, it will continue to benefit from its simplicity. As remarked in Section 2.1, in practice the amount of time and the cost of improving the procedures have to be related to the benefits and, hence, an approach which is roughly right might very well be preferable.

In particular the implementation of the standardized procedures suggested by Pinheiro *et al.* (2003) and in Paper I is much more troublesome than the implementation of England's method. Hence, arguing that the procedures in Paper I are theoretically more correct than England's method, how much would we benefit from using them? Would it be worth spending time on the implementation?

Table 4.1 shows a comparison between England's method and the bootstrap procedures suggested in Paper I for the data set provided in Taylor & Ashe (1983), which has been used throughout all papers. Here, 10 000 iterations were used for each approach. As described in Section 2.5, the process error of England's approach is included by a second sampling from a full distribution conditionally on the first bootstrap sampling. England uses either an ODP distribution parameterized by the mean $\hat{m}^*_{ij}$ or a gamma distribution with mean $\hat{m}^*_{ij}$ and variance $\hat{\phi}\hat{m}^*_{ij}$. Note that this variance assumption differs from the assumption of $p = 2$ in (2.17) and the corresponding full gamma distribution with mean $\hat{m}_{ij}$ and variance $\hat{\phi}\hat{m}^2_{ij}$ for the non-parametric and the parametric bootstrap procedures, respectively, in Paper I. England's choice of variance function can still be motivated by the observation that the two first moments of an ODP distribution often fits the data quite well, while the ODP distribution itself is an unrealistic distribution due to its support. Anyway, it is not relevant to compare England's approach with the procedures in Paper I under the assumption of $p = 2$ or the corresponding gamma distribution.

Recall that the procedures suggested in Paper I adopt either standardized or unstandardized prediction errors and, hence, are referred to as standardized or unstandardized bootstrap procedures. Moreover, the two residuals that equal zero in the south and the east corner of the triangle have been removed for England's method as well as the two non-parametric procedures in Paper I, see e.g. England (2002) for details. Also note that the results of England's method do not perfectly coincide with the results presented in England (2002). Some reasons might be that the residuals equaling zero have been removed, a larger number of iterations has been used and the algorithms adopted for the generation of random numbers might slightly differ for the modeling softwares.

As can be seen in Table 4.1, all methods result in approximately the same prediction error, i.e. the standard deviation of the bootstrap samples. For the 95th percentile, the procedures suggested in Paper I result in a $2-5\%$ lower estimate compared to England's method, while the difference is larger for the 99.5th percentile where a $5-11\%$ lower value is obtained.

For the unstandardized bootstrap, the difference can partly be explained by the opposite shift of the bootstrap means relatively the reserve estimate. Interestingly, the difference between the estimated reserve and the bootstrap mean is reduced by the standardization. The positive shift of England's method is usually adjusted for practical applications, which then yields lower estimates of the upper limits.

Let us now return to the question of whether it is worth implementing a possibly more troublesome procedure instead of using England's method. If the goal would be to estimate the prediction error, then we will obviously not ben-

	England ODP	England Gamma	Unstand. Non-par. $p=1$	Unstand. Par. ODP	Stand. Non-par. $p=1$	Stand. Par. ODP
Est. res.	18 680 856	18 680 856	18 680 856	18 680 856	18 680 856	18 680 856
Mean	19 002 077	19 032 480	18 525 343	18 553 504	18 738 917	18 729 533
Mean−Est. res.	321 221	351 624	-155 513	-127 352	58 061	48 677
Stand. dev.	3 028 383	3 034 122	3 053 277	3 006 016	2 924 040	3 018 143
95th perc.	24 301 829	24 305 631	23 189 995	23 178 763	23 528 736	23 804 581
99.5th perc.	28 352 134	28 156 002	25 467 162	25 316 634	26 589 078	26 750 057

Table 4.1: Bootstrap statistics for England's method, using either an ODP or a gamma distrubution for the process error, and the corresponding procedures suggested in Paper I when $p = 1$ and an ODP distribution is assumed for the non-parametric and the parametric bootstrap procedure, respectively. The data set provided in Taylor & Ashe (1983) is used.

efit from choosing a method which requires more implementation time. The conslusion might be the same if we for some reason would like to estimate the 95th percentile. However, suppose that the insurance company is required to hold risk capital corresponding to the 99.5th percentile for all reserving classes and the results above are systematic. A switch of methods could then result in a 10% decrease of the required capital, which probably would be considered as a very good argument in order to spend time on the implementation of a new, possibly more troublesome, method.

The point to be made regarding the example above is that even though England's method was questioned in Paper I it might very well be preferable for real applications in practice - it just depends on the task at hand.

4.2 The standardized versus the unstandardized bootstrap

In Paper I it is argued that the standardized prediction error

$$\text{pe}^{**} = \frac{R^{**} - \hat{R}^{*}}{\sqrt{\widehat{Var}(R^{**})}} \tag{4.1}$$

used in Pinheiro *et al.* (2003) increases the accuracy of the simulated predictive distribution compared to the alternative unstandardized prediction error

$$\text{pe}^{**} = R^{**} - \hat{R}^{*}\,. \tag{4.2}$$

Hence, rather than assuming that $R - \hat{R}$ and $R^{**} - \hat{R}^{*}$ have similar distributions the standardized approach instead assumes that $(R - \hat{R})/\sqrt{\widehat{Var}(R)}$ and $(R^{**} - \hat{R}^{*})/\sqrt{\widehat{Var}(R^{**})}$ have similar distributions.

Many references comment on this topic, see, for example, Davison & Hinkley (1997), Garthwaite *et al.* (2002), Hall (1995) and Hjorth (1994). Garthwaite *et al.* (2002) use a quantity which would correspond to

$$\text{pe}^{**} = \frac{R^{**} - \hat{R}^{*}}{\sqrt{\widehat{Var}(\hat{R}^{*})}}\,, \tag{4.3}$$

in contrast to (4.1), and refer to the approach as the *bootstrap t method*. Moreover, they remark that the research literature suggests that the use of (4.3) increases the accuracy in the estimation of the upper limits of a confidence interval compared to (4.2). The inaccuracy of the estimated upper limits is of order $O(n^{-\frac{3}{2}})$ for the suggested standardized approach, while it is of order $O(n^{-1})$ for the unstandardized approach. Recall that n is the number of observations and pseduo-observations in the original triangle and the simulated triangles, respectively.

Note that the order of inaccuracy of the estimated upper limits holds for estimation problems, while we are focusing on a prediction problem. Moreover, as remarked in Section 3.1 it is by no means clear that large triangles would yield an increased accuracy of the estimates, since parameters corresponding to late origin and development years are hard to estimate for large as well as small development triangles.

Instead of (4.1), Paper I suggests the use of

$$\mathrm{pe}^{**} = \frac{R^{**} - \hat{R}^{*}}{\sqrt{\widehat{Var}(R^{**} - \hat{R}^{*})}}, \tag{4.4}$$

since the estimation error tend to be larger than the process error. Moreover, it is suggested that $\sqrt{\widehat{Var}(R^{**} - \hat{R}^{*})}$ could be achieved by means of a double bootstrap. Garthwaite *et al.* (2002) also mention this alternative in order to obtain the denominator of (4.3), however, they refer to it as *second-level bootstrapping*. They also suggest the use of the jackknife for this purpose.

The standardized and the unstandardized bootstrap procedures are numerically compared in Paper I and the conclusion, which can also be seen in Table 4.1, is that the standardized approach yields a larger estimate of the upper limits than the unstandardized approach. This was explained by the left skewness of the predictive distribution of the unstandardized bootstrap compared to the distribution obtained by the standardized bootstrap.

The double bootstrap was not numerically investigated in Paper I due to its computational complexity. The nested sampling loop implies very long running times if an ordinary PC is used for more than 1 000 (double) iterations. Since the estimation of the denominator of (4.4) is robust even for a relatively small number of iterations, an alternative might be to use, say, 1 000 iterations for the second-level bootstrap, while keep 10 000 iterations for the first-level bootstrap.

Let B_1 and B_2 denote the number of iterations in the first-level and second-level bootstrap, respectively. Table 4.2 presents results of the double parametric bootstrap, under the assumption of an ODP, for different choices of B_1 and B_2 (limited by the running time). As we can see, the results seem quite robust even for a low number of iterations. Moreover, the upper limits of the double bootstrap are larger than for the corresponding standardized bootstrap in Table 4.1, since the means have been shifted towards right.

Note that the approach in England (2002) resembles the unstandardized method even though it is another approach. Interestingly, the double bootstrap, which is supposed to be an improvement of the standardized bootstrap which, in turn, is supposed to be an improvement of the unstandardized bootstrap, yields results which are similar to the ones obtained by England's method. Again, if these results are systematic and, hence, England's approach is a fair approximation, then this method is certainly preferable since it is easier to implement and much faster to run.

	$B_1 = 500$ $B_2 = 500$	$B_1 = 1\,000$ $B_2 = 1\,000$	$B_1 = 3\,000$ $B_2 = 3\,000$	$B_1 = 10\,000$ $B_2 = 500$	$B_1 = 10\,000$ $B_2 = 1\,000$
Est. res.	18 680 856	18 680 856	18 680 856	18 680 856	18 680 856
Mean	18 987 008	18 974 593	19 011 506	18 924 153	18 956 413
Mean−Est. res.	306 152	293 737	330 650	243 297	275 557
Stand. dev.	3 019 459	3 019 459	2 927 042	3 078 212	3 093 998
95th perc.	24 128 488	24 623 449	24 132 185	24 339 556	24 578 003
99.5th perc.	29 084 282	28 512 489	27 943 754	28 285 594	28 669 889

Table 4.2: Bootstrap statistics for the double parametric bootstrap when an ODP is assumed. Here, B_1 and B_2 denote the number of iterations in the first-level and second-level bootstrap, respectively. The data set provided in Taylor & Ashe (1983) is used.

To this, it would be interesting to conduct an extensive investigation based on simulated (for which the true reserves are known) as well as real data, to compare England's approach with the double bootstrap and other standardized and unstandardized bootstrap approaches described above.

4.3 Prediction versus estimation

The assumption of an underlying GLM is frequently adopted in stochastic claims reserving. The model parameters are obtained by maximum likelihood estimation, while the model structure is used for prediction of future outcomes rather than estimation of quantities related to the observations. Hence, applying bootstrap methods to claims reserving often implies that they are used in order to obtain predictive intervals for GLMs.

It would of course be comforting to be able to base the application on theoretical research, however, even though the bootstrap in general has been thoroughly explored in the literature, little has been said regarding its application to prediction problems for GLMs.

Another approach is to combine GLMs with Bayesian sampling methods. In England & Verrall (2006) serveral Bayesian methods are suggested for claims reserving purposes and they are numerically studied using the data set from Taylor & Ashe (1983). Consequently, they could be used for benchmarking and a comparison shows that the results of the Bayesian approaches are similar to the ones presented here.

4.4 Research versus practical applications

As indicated in Section 2.2, the reserving methods suggested in the research literature seldom fit practical applications perfectly and, hence, the actuaries often have to make ad hoc adjustments. Moreover, in the literature the suggested methods tend to be illustrated using neat and uncomplicated data sets and are therefore not so often questioned.

Even though this thesis aims at providing practical solutions to practical problems it is still no exception from the statement above. As soon as a practical problem is adopted as a research problem there is always a risk that it looses some of its practical appeal. One reason is that the theoretical level of the presentation often has to be increased in order to get the results published in scientific journals. Moreover, there is unfortunately little scientific interest in reproducing the studies more than a few times on other data sets in order to discover real problems and provide solutions to them.

Paper I and Paper III clearly demonstrate the difference between research and practical applications since the arguments of the two papers are pretty much each other's opposites. In Paper I we adopt the practitioner's point of view arguing that the stochastic models suggested in the literature are discordant with the actuary's way of working. Therefore we relax the model assumptions in order to develop a more flexible bootstrap procedure that suits our practical needs. However, in Paper III we adopt the researcher's point of view arguing that the actuary's way of working is somewhat ad hoc. Thus, we assume a more detailed stochastic model in order to systematize a reserving exercise which includes smoothing of development factors. This approach is theoretically more correct, but less flexible and perhaps less useful for practical applications.

5. Reserve risk in a business model

5.1 Solvency II

So far the insurance business as well as the authorities' supervision have been based on a general conservativeness regarding the liabilities to the policy holders. There are laws that dictate how much capital the firms must hold and how it may be invested, see Försäkringsrörelselagen by Sveriges Riksdag (1982) for the regulations applied in Sweden today. However, the current regulations rather consider the volume than the risk of the business in the calculation of the required amount of capital.

In order to capture the individual characteristics of the firms the regulations are being modernized within EU. According to the Solvency II Directive 2009/138/EC by the European Parliament and the Council of the European Union (2009), the required capital will instead be calculated by quantifying the risks of the firm under market-like assumptions. The authorities will provide a standard formula which consider the major risks that an insurance company is exposed to, but own internal models will also be allowed. For instance, the firms will have to quantify premium and reserve risk, catastrophe risk, market risks such as e.g. equity risk, interest rate risk and currency risk, counterparty default risk and operational risk. For Solvency II purposes the internal models will have to be stochastic, a one-year time perspective should be adopted and the risks should be measured according to a 99.5% quantile. Furthermore, the purpose of an internal model is not only to be a supervisory tool - it has to be used in the business as well in order to show its trustworthiness. Potential areas of use could be e.g. business planning, investment strategies, purchase of reinsurance and pricing.

The analysis of the business by such an internal simulation model is often referred to as Dynamic Financial Analysis (DFA) in non-life insurance. Kaufmann *et al.* (2001) gives an introduction to DFA and also provides an example of a model.

5.2 The one-year reserve risk

Thus, for Solvency II purposes the amount of capital that the insurance company must hold in order to be able to handle a negative run-off result the next accounting year with 99.5% probability is of interest. The one-year run-off result is defined as the difference between the opening reserve at the beginning of the year and the sum of payments during the year and the closing reserve of the same portfolio at the end of the year. Thus, if we at the end of year t want to make predictions of the run-off result at the end of the unobserved year $t+1$, and if we do not add neither a new accident year nor a new development year, we have to find the predictive distribution of

$$\hat{R}^t - (\sum_{i=1}^{t} C_{i,t+1-i} + \hat{R}^{t+1}), \tag{5.1}$$

where $\hat{R}^t$ and $\hat{R}^{t+1}$ are the estimated reserves at the end of year t and $t+1$, respectively.

Ohlsson & Lauzeningks (2009) provide details for how the one-year reserve risk could be obtained by bootstrapping. In order to implement such a procedure here, $\hat{R}^t$ is assumed to be obtained by the chain-ladder method. The claims paid during the year, called $\tilde{C}_{i,t+1-i}$ for $i = 1,\ldots,t$, are simulated by bootstrapping. Hence, B new triangles, corresponding to the potential outcome of year $t+1$, can be obtained by adding the simulated diagonals of paid amounts to the original triangle of year t. The chain-ladder method is then applied to each of the new triangles resulting in B simulated values of $\tilde{R}^{t+1}$. Finally, the B values of

$$\hat{R}^t - (\sum_{i=1}^{t} \tilde{C}_{i,t+1-i} + \tilde{R}^{t+1}) \tag{5.2}$$

are analyzed to obtain an estimate of the 0.5th percentile corresponding to the capital that the insurance company must hold in order to be able to deal with a negative run-off result with 99.5% probability. Let this procedure be denoted by 'Method I'.

Now suppose that the predictive distribution of $\sum_{i=1}^{t} C_{i,t+1-i} + \hat{R}^{t+1}$ in (5.1) is approximated by the predictive distribution of $\hat{R}^t$ obtained by bootstrapping. This assumption implies that we do not adopt the re-reserving procedure described above, but assume the same relative uncertainty in $\hat{R}^{t+1}$ as in $\hat{R}^t$ even though new information will be available at year $t+1$. Let the approach implemented under this assumption be denoted by 'Method II'.

The modeling of the quantity in (5.1) could be simplified further. Assume that $\sum_{i=1}^{t} C_{i,t+1-i} + \hat{R}^{t+1}$ is either normal, lognormal or gamma distributed and let these distributions be parameterized by the mean $\hat{R}_t$ and the standard deviation corresponding to the prediction error of $\hat{R}_t$. Let these approaches be denoted by 'Method III Normal', 'Method III Lognormal' and 'Method III Gamma', respectively.

Table 5.1 presents statistics for the run-off result of the data set in Taylor & Ashe (1983) when Methods I-III have beeen implemented using England's bootstrap method. Here $B = 10\,000$ iterations are used.

As can be seen, the insurance company will benefit from using the more precise Method I for the portfolio under investigation since the 0.5th percentile is reduced by about 20% compared to the approximation adopted in Method II. Method III Normal yields almost the same capital requirement as Method I even though the shape of the distribution is different. In order to be allowed to use this simplified approach the insurance company would have to explain why the normal distribution would be a proper assumption.

Note that the simulation means of the three methods differ. The distribution of the run-off result obtained by Method III is centered around origin, while it is shifted for Method I and Method II. Moreover, note that for short tailed portfolios the difference in the results of Methods I-II is decreased since the one-year risk approaches the ultimate risk.

	Method I	Method II	Method III Normal	Method III Lognormal	Method III Gamma
$\hat{R}_t$	18 680 856	18 680 856	18 680 856	18 680 856	18 680 856
Mean	-204 170	-374 488	21 130	-20 560	43 807
Stand. dev.	2 475 731	3 050 364	3 091 834	3 108 941	3 062 081
5th perc.	-4 601 721	-5 778 777	-5 048 851	-5 501 742	-5 336 509
0.5th perc.	-7 777 637	-9 782 293	-7 951 939	-9 463 218	-8 507 867

Table 5.1: Statistics for the run-off result of the data set in Taylor & Ashe (1983).

6. Some corrections

Paper I, Appendix A, page 24: Here t is denoting the development years as well as the variable in the moment generating function.
Paper I, Appendix B, pages 25-26: The equations have been labeled by (B.1)-(B.6) but are referred to by (A.1)-(A.6).

7. References

Davison, A. C. & Hinkley, D. V. (1997): Bootstrap Methods and their Applications. *Cambridge University Press*.

Efron, B. & Tibshirani, R.J. (1993): An Introduction to the Bootstrap. *Chapman & Hall*, N. Y.

England, P. (2002): Addendum to “Analytic and Bootstrap Estimates of Prediction Error in Claims Reserving”. *Insurance: Mathematics and Economics*, **31**, 461-466.

England, P. & Verrall, R. (2001): A Flexible Framework for Stochastic Claims Reserving. *Proceedings of the Casualty Actuarial Society*, **LXXXVIII**, 1-38.

England, P. & Verrall, R. (1999): Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving. *Insurance: Mathematics and Economics*, **25**, 281-293.

England, P. & Verrall, R. (2006): Predictive Distributions of Outstanding Liabilities in general Insurance. *Annals of Actuarial Science*, **1**, (2), 221-270.

England, P. & Verrall, R. (2002): Stochastic Claims Reserving in General Insurance. *British Actuarial Journal*, **8** (3), 443-518.

European Parliament and the Council of the European Union (2009): Directive 2009/138/EC. Available online at: http://eur-lex.europa.eu

Faculty and Institute of Actuaries (1997): Claims Reserving Manual. http://www.actuaries.org.uk/research-and-resources/pages/library-services

Garthwaite, P., Jolliffe, I. & Jones, B. (2002): Statistical Inference. *Oxford University Press*.

Hall, P. (1995): The Bootstrap and Edgeworth Expansion. *Springer*.

Hjort, U. (1994): Computer Intensive Statistical Methods. *Chapman & Hall*.

Kaufmann, R., Gadmer, A. & Klett, R. (2001): Introduction to Dynamic Financial Analysis. *ASTIN Bulletin*, 31(1), 213-249.

Mack, T. (1993): Distribution Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *ASTIN Bulletin*, **23**, 213-225.

Mack, T. & Venter, G. (2000): A Comparison of Stochastic Models that Reproduce Chain Ladder Reserve Estimates. *Insurance: Mathematics and Economics*, **26**, 101-107.

Ohlsson, E. & Lauzeningks, J. (2009): The One-Year Non-Life Insurance Risk. *Insurance: Mathematics and Economics*, **45**, (2), 203-208.

Pinheiro, P.J.R., Andrade e Silva, J.M. & Centeno, M.d.L (2003): Bootstrap Methodology in Claim Reserving. *The Journal of Risk and Insurance*, **4**, 701-714.

Renshaw, A. (1994): Modelling the Claims Process in the Presence of Covariates. *ASTIN Bulletin*, **24**, 265-286.
Renshaw, A. & Verrall, R. (1998): A Stochastic Model Underlying the Chain-Ladder Technique. *British Actuarial Journal*, **4**, IV, 903-923.
Sveriges Riksdag (1982): Försäkringsrörelselagen 1982:713. Available online at: http://www.riksdagen.se
Taylor, G. (2000): Loss Reserving-An Actuarial Perspective. *Boston: Kluwer Academic Press.*
Taylor, G. (1977): Separation of Inflation and Other Effects from the Distribution of Non-life Insurance Claims Delays. *ASTIN Bulletin*, **9**, 217-230.
Taylor, G. & Ashe, F.R. (1983): Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics*, 23, 37-61.
Teugels, J. L. & Sundt, B. (2004): Encyclopedia of Actuarial Science. *Wiley*.
Verbeek, H. G. (1972): An Approach to the Analysis of Claims Experience in Motor Liability Excess of Loss Reassurance. *ASTIN Bulletin*, **6**, 195-202.
Verrall, R. (2000): An Investigation into Stochastic Claims Reserving Models and the Chain-ladder Technique. *Insurance: Mathematics and Economics*, **26**, 91-99.
Verrall, R. (1991): On the Estimation of Reserves from Loglinear Models. *Insurance: Mathematics and Economics*, **10**, 75-80.
Verrall, R. & England, P. (2000): Comments on: "A Comparison of Stochastic Models that reproduce Chain Ladder Reserve Estimates, by Mack and Venter." *Insurance: Mathematics and Economics*, **26**, 109-111.
Wütrich, M. & Merz, M. (2008): Stochastic Claims Reserving Methods in Insurance. *John Wiley & Sons Ltd*.
Zehnwirth, B. (1989): Regression methods - Applications. *Conference paper, 1989 Casualty Loss Reserve Seminar*. Available online at: http://www.casact.org/pubs/CLRSTrans/1989/825.pdf.