BOOTSTRAPPING THE SEPARATION METHOD IN CLAIMS RESERVING

BY

SUSANNA BJÖRKWALL, OLA HÖSSJER AND ESBJÖRN OHLSSON

ABSTRACT

The separation method was introduced by Verbeek (1972) in order to forecast numbers of excess claims and it was developed further by Taylor (1977) to be applicable to the average claim cost. The separation method differs from the chain-ladder in that when the chain-ladder only assumes claim proportionality between the development years, the separation method also separates the claim delay distribution from influences affecting the calendar years, e.g. inflation. Since the inflation contributes to the uncertainty in the estimate of the claims reserve it is important to consider its impact in the context of risk management, too.

In this paper we present a method for assessing the prediction error distribution of the separation method. To this end we introduce a parametric framework within the separation model which enables joint resampling of claim counts and claim amounts. As a result, the variability of Taylor's predicted reserves can be assessed by extending the parametric bootstrap techniques of Björkwall *et al.* (2009). The performance of the bootstrapped separation method and chain-ladder is compared for a real data set.

KEYWORDS

Bootstrap; Chain-ladder; Development triangle; Inflation; Separation method; Stochastic claims reserving.

1. INTRODUCTION

One issue for the reserving actuary is how to deal with inflation, which contributes to the uncertainty in the estimate of the claims reserve. Even though some proper reserving techniques are suggested in the literature, little has been said about how to approach this issue when it comes to finding the variability of the actuary's best estimate either analytically or by bootstrapping.

Due to external forces the average cost per claim will change from one calendar year to another. Typically this *claims inflation* is specific to each line of business and depends on the economic inflation, which usually can be tied

Astin Bulletin 40(2), 845-869. doi: 10.2143/AST.40.2.2061138

to some relevant index, as well as on factors like legislation and attitudes to policy holder compensation. The latter result in so called *superimposed claims inflation*.

The chain-ladder method makes implicit allowance for claims inflation since it projects the inflation present in the past data into the future, see e.g. Taylor (2000). Consequently, it only works properly when the inflation rate is constant. When the economic inflation rate is non-constant, the past paid losses can be converted to current value by some relevant index before they are projected into the future by the chain-ladder, but still there is no allowance for superimposed claims inflation.

Another approach of dealing with claims inflation is to incorporate it into the model underlying the reserving method. In this way the past inflation rate can be estimated and the future inflation rate can be predicted within the model. Verbeek (1972) introduced such a method in the reinsurance context and Taylor (1977) developed it further to be applicable to the average claim cost in a general context. The reserving technique is called *the separation method*. However, the separation method has, unlike its famous relative, remained quite anonymous in the literature on stochastic claims reserving. For instance, the mean squared error of prediction (MSEP) for the chain-ladder was analytically calculated by Mack (1993) and revisited by Buchwalder *et al.* (2006) and Mack *et al.* (2006) and a full predictive distribution was obtained for the chain-ladder by bootstrapping in England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003). Recently the variability of other reserving methods has been investigated as well, e.g. the Bornhuetter-Ferguson method by analytical approximation in Mack (2008) and the Munich chain-ladder, see Quarg & Mack (2004), by bootstrapping of two correlated quantities in Liu & Verrall (2008).

The object of this paper is to analyze the variability of the separation method. Since bootstrapping easily gives a full predictive distribution and can also be used in risk management with Dynamic Financial Analysis (DFA) we develop a bootstrap procedure for the separation method. For this purpose we use an extended version of the parametric bootstrap technique described in Björkwall *et al.* (2009). To this end, we introduce a parametric framework within the separation model, in which 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.

Section 2 contains the definitions and the theory behind the separation method. In Section 3 the suggested bootstrap methodology is presented and it is studied numerically for the well-known data set from Taylor & Ashe (1983) in Section 4. Finally, Section 5 contains a discussion regarding the chosen approach.

2. The separation method

Assume that we have a triangle of incremental observations of paid claims $\{C_{ij}; i, j \in \nabla\}$, where ∇ denotes the upper, observational triangle $\nabla = \{i = 0, \ldots, t;$

$j = 0, \ldots, t - i\}$. The suffixes i and j refer to the origin year and the development year, respectively, see Table 2.1. In addition, the suffix $k = i + j$ is used for the calendar years, i.e. the diagonals of ∇. The purpose 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. Furthermore, assume that we have a triangle of the incremental observations of the number 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 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 and Δ_i denotes the rows corresponding to origin year i in the upper triangle ∇ and the lower triangle Δ, respectively.

TABLE 2.1

THE TRIANGLE ∇ OF OBSERVED INCREMENTAL PAYMENTS

	Development year					
Accident year	0	1	2	⋯	$t-1$	t
0	C_{00}	C_{01}	C_{02}	⋯	$C_{0,t-1}$	$C_{0,t}$
1	C_{10}	C_{11}	C_{12}	⋯	$C_{1,t-1}$	
2	C_{20}	C_{21}	C_{22}	⋯		
⋮	⋮	⋮	⋮			
$t-1$	$C_{t-1,0}$	$C_{t-1,1}$				
t	$C_{t,0}$					

TABLE 2.2

THE TRIANGLE Δ OF UNOBSERVED FUTURE CLAIM COSTS

	Development year					
Accident year	0	1	2	⋯	$t-1$	t
0						
1						$C_{1,t}$
2					$C_{2,t-1}$	$C_{2,t}$
⋮					⋮	⋮
$t-1$			$C_{t-1,2}$	⋯	$C_{t-1,t-1}$	$C_{t-1,t}$
t		$C_{t,1}$	$C_{t,2}$	⋯	$C_{t,t-1}$	$C_{t,t}$

The separation method is based on the assumption that N_i is considered as known. Since the number of claims is often finalized quite early even for long-tailed business, N_i may very well be estimated separately, e.g. by the chain-ladder if a triangle of claim counts is provided, and then be treated as known. Henceforth estimates $\hat{n}_{ij}$ of the expectations $n_{ij} = E(N_{ij})$ are obtained by the chain-ladder for all cells in both ∇ and Δ. The estimated ultimate number of claims relating to origin year i is then

$$\hat{N}_i = \sum_{j \in \nabla_i} N_{ij} + \sum_{j \in \Delta_i} \hat{n}_{ij}. \tag{2.2}$$

The chain-ladder method operates on cumulative claim counts

$$A_{ij} = \sum_{\ell=0}^{j} N_{i\ell} \tag{2.3}$$

rather than incremental claim counts N_{ij}. Let $v_{ij} = E(A_{ij})$. Development factors g_j are estimated for $j = 0, 1, \ldots, t-1$ by

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

yielding

$$\hat{v}_{ij} = A_{i,t-i}\ \hat{g}_{t-i}\ \hat{g}_{t-i+1} \cdots \hat{g}_{j-1} \tag{2.5}$$

and

$$\hat{n}_{i,j} = \hat{v}_{i,j} - \hat{v}_{i,j-1} \tag{2.6}$$

TABLE 2.3

THE TRIANGLE ∇ OF OBSERVED INCREMENTAL NUMBERS OF REPORTED CLAIMS

	Development year					
Accident year	0	1	2	⋯	$t-1$	t
0	N_{00}	N_{01}	N_{02}	⋯	$N_{0,t-1}$	$N_{0,t}$
1	N_{10}	N_{11}	N_{12}	⋯	$N_{1,t-1}$	
2	N_{20}	N_{21}	N_{22}	⋯		
⋮	⋮	⋮	⋮			
$t-1$	$N_{t-1,0}$	$N_{t-1,1}$				
t	$N_{t,0}$					

for Δ, while estimates of $\hat{v}_{ij}$ for ∇ are obtained by the process of backwards recursion described in England & Verrall (1999).

While the chain-ladder only assumes claim proportionality between the development years, the separation method in Taylor (1977) separates the claim delay distribution from influences effecting the calendar years, e.g. inflation. In the separation model we first assume that the proportion of the average claim amount paid in development year j is constant over i; denote this proportion by r_j. If the claims are fully paid by year t we have the constraint

$$\sum_{j=0}^{t} r_j = 1. \tag{2.7}$$

We then make a further assumption that the claim amount is proportional to some index, say λ_k, that relates to the calendar year k during which the claims are paid. The expected average claim cost for development year j and calendar year k is then proportional to $r_j \lambda_k$.

The separation model can be given the following formulation, which is at a bit more detailed level than the one given in Taylor (1977). Let C_{ijl} denote the amount paid during calendar year k for the l:th individual claim incurred in origin year i and assume that C_{ijl} are conditionally independent for all i, j and l given N_i. According to the discussion above we also assume that

$$E(C_{ijl} \,|\, N_i) = r_j \lambda_k. \tag{2.8}$$

Since the total amount paid during calendar year k for claims incurred in origin year i is

$$C_{ij} = \sum_{l=1}^{N_i} C_{ijl} \tag{2.9}$$

TABLE 2.4

THE TRIANGLE ∇ OF EXPECTED AVERAGE CLAIM COST

	Development year					
Accident year	0	1	2	⋯	$t-1$	t
0	$r_0\lambda_0$	$r_1\lambda_1$	$r_2\lambda_2$	⋯	$r_{t-1}\lambda_{t-1}$	$r_t\lambda_t$
1	$r_0\lambda_1$	$r_1\lambda_2$	$r_2\lambda_3$	⋯	$r_{t-1}\lambda_t$	
2	$r_0\lambda_2$	$r_1\lambda_3$	$r_2\lambda_4$	⋯		
⋮	⋮	⋮	⋮			
$t-1$	$r_0\lambda_{t-1}$	$r_1\lambda_t$				
t	$r_0\lambda_t$					

we obtain

$$E\left(\frac{C_{ij}}{N_i}\middle| N_i\right) = \frac{1}{N_i}\sum_{l=1}^{N_i} E\big(C_{ijl}\,|\,N_i\big) = \frac{1}{N_i}\sum_{l=1}^{N_i} r_j\,\lambda_k = r_j\,\lambda_k \tag{2.10}$$

for the conditional expectation of the average claim costs given the ultimate number of claims and this relation is the basic assumption of the separation method. The expectations in equation (2.10) now build up the triangle in Table 2.4.

If N_i is estimated separately by (2.2), it follows from (2.8) and (2.9) that

$$\begin{aligned} E\left(\frac{C_{ij}}{\hat{N}_i}\middle|\nabla N\right) &= \frac{E(E(C_{ij}\,|\,N_i,\nabla N)\,|\,\nabla N)}{\hat{N}_i} \\ &= r_j\,\lambda_k\,\frac{\left(\sum_{\nabla_i} N_{ij} + \sum_{\Delta_i} n_{ij}\right)}{\left(\sum_{\nabla_i} N_{ij} + \sum_{\Delta_i} \hat{n}_{ij}\right)} \\ &\approx r_j\,\lambda_k \end{aligned} \tag{2.11}$$

where in the last equality we used $\hat{n}_{ij} \approx n_{ij}$.

Estimates $\hat{r}_j$ and $\hat{\lambda}_k$ of the parameters in the triangle in Table 2.4 can now be obtained by marginal sum estimation using the corresponding triangle ∇s of observed values

$$s_{ij} = \frac{C_{ij}}{\hat{N}_i}, \tag{2.12}$$

and the marginal sum equations

$$s_{k0} + s_{k-1,1} + \ldots + s_{0k} = \big(\hat{r}_0 + \ldots + \hat{r}_k\big)\hat{\lambda}_k, \quad k = 0, \ldots, t \tag{2.13}$$

for the diagonals of ∇ and

$$s_{0j} + s_{1j} + \ldots + s_{t-j,j} = \left(\hat{\lambda}_j + \ldots + \hat{\lambda}_t\right)\hat{r}_j, \quad j = 0, \ldots, t \tag{2.14}$$

for the columns of ∇.

Taylor (1977) shows that the equations (2.13) - (2.14), with the side constraint (2.7), have a unique solution that can be obtained recursively, starting with $k = t$ in (2.13) to solve for $\hat{\lambda}_t$, then $j = t$ in (2.14) to solve for $\hat{r}_t$, $k = t-1$ in (2.13) to solve for $\hat{\lambda}_{t-1}$ and so on. This yields

$$\hat{\lambda}_k = \frac{\sum_{i=0}^{k} s_{i,k-i}}{1 - \sum_{j=k+1}^{t} \hat{r}_j}, \quad k = 0, \ldots, t \tag{2.15}$$

$$\hat{r}_j = \frac{\sum_{i=0}^{t-j} s_{ij}}{\sum_{k=j}^{t} \hat{\lambda}_k}, \quad j = 0, \ldots, t, \tag{2.16}$$

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

Estimates $\hat{m}_{ij}$ of the expectations $m_{ij} = E(C_{ij})$ for cells in ∇ are now given by

$$\hat{m}_{ij} = \hat{N}_i \hat{r}_j \hat{\lambda}_k, \tag{2.17}$$

but in order to obtain the estimates of Δ it remains to predict λ_k for $t + 1 \le k \le 2t$, which requires some inflation assumption.

If there is a trend in the inflation indexes λ_k for $k \le t$ then smoothing and extrapolation could be used in order to forecast the future inflation. An alternative is to use an average of the past indexes. In any case, with an inflation assumption of, say, K%, the forecasted λ_{k+1} can be obtained as

$$\hat{\lambda}_{k+1} = \left(1 + \frac{K}{100}\right) \hat{\lambda}_k, \quad t \le k \le 2t - 1. \tag{2.18}$$

The cell expectations of ΔC_{ij} are then estimated by equation (2.17) and estimators of the outstanding claims are obtained by summing per accident year $\hat{R}_i = \sum_{j \in \Delta_i} \hat{m}_{ij}$. The estimator of the total reserve is $\hat{R} = \sum_{\Delta} \hat{m}_{ij}$.

The separation model described by Taylor (1977) is more general than the one discussed in this paper, since the original model does not presume that N_i is the number of claims; it could be some other exposure relating to origin year i as well. However, in this paper we stick to the number of claims.

3. A CONDITIONAL PARAMETRIC BOOTSTRAP APPROACH

For the purpose of obtaining the predictive distribution of the claims reserve R estimated by the separation method the bootstrap technique described in Pinheiro *et al.* (2003) and, in particular, the parametric approach in Björkwall *et al.* (2009) is used. For the sampling process we model the paid claims C_{ij} conditionally on N_i in accordance with (2.11). We provide models for the assumption of stochastic N_i as well as for the case when N_i is considered as known. The former assumption demands that we develop the technique described in Björkwall *et al.* (2009) in order to handle ∇N as well as ∇C.

3.1. Stochastic Poisson distributed claim counts

Verbeek (1972) adopted a Poisson distribution for the claim counting variable, while the method described in Taylor (1977) is distribution-free. However, the assumption of independent and Poisson distributed claim counts

$$N_{ij} \in Po(n_{ij}) \tag{3.1}$$

yields a very reasonable model for the sampling process.

In addition we assume that the conditionally independent claims $C_{ijl} \,|\, N_i$ in (2.8) are gamma distributed. We use the notation

$$C_{ijl} \,|\, N_i \in \Gamma\Big(\frac{1}{\phi}, r_j \lambda_k \phi\Big), \tag{3.2}$$

where $1/\phi$ is the so called index parameter and $r_j \lambda_k \phi$ is the scale, so that the expected value is $r_j \lambda_k$. Moreover, $\phi > 0$.

Recalling (2.9) and the independence of the C_{ijl} we find that

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

which is consistent with (2.10) since

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

The variance of the amounts in (3.3) is

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

which corresponds to a weighted generalized linear model under the assumption of a logarithmic link function and a gamma distribution. We use a Pearson type estimate of ϕ, cf. McCullagh & Nelder (1989),

$$\hat{\phi} = \frac{1}{|\nabla| - q} \sum_{\nabla} \hat{N}_i \frac{\big(C_{ij} - \hat{E}(C_{ij} \,|\, N_i)\big)^2}{\hat{E}^2(C_{ij} \,|\, N_i)} = \frac{1}{|\nabla| - q} \sum_{\nabla} \hat{N}_i \frac{\big(C_{ij} - \hat{N}_i \hat{r}_j \hat{\lambda}_k\big)^2}{\big(\hat{N}_i \hat{r}_j \hat{\lambda}_k\big)^2}, \tag{3.6}$$

where $|\nabla| = (t+1)(t+2)/2$ is the number of observations in ∇C, the estimators $\hat{N}_i$, $\hat{\lambda}_j$ and $\hat{r}_j$ are obtained from (2.2), (2.15) and (2.16) and $q = 2t + 1$ is the number of parameters that have to be estimated by the separation method, i.e. r_j for $j = 0, 1, \ldots, t-1$ and λ_k for $k = 0, 1, \ldots, t$.

Note that (3.3) could be interpreted as follows if the inflation is constant, i.e. $\lambda_k = \lambda$; given N_i claims we allocate claim amounts independently over the development years j with amounts whose means are proportional to $r_0, \ldots, r_t$. According to (3.2) we not only allocate claim amounts but individual claims as well. This interpretation is consistent with the assumptions discussed in Section 2.

We adopt the bootstrap technique described in Pinheiro *et al.* (2003) and, in particular, the parametric approach in Björkwall *et al.* (2009). 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. Hence, the process error is included in R^{**}, i.e. the true outstanding claims in the bootstrap world, while the estimation error is included in $\hat{R}^*$, i.e. the estimated outstanding claims in the bootstrap world. Henceforth we use the index * for random variables or plug-in estimators in the bootstrap world which correspond to observations or estimators in the real world, while the index ** is used for random variables in the bootstrap world when the counterparts in the real world are unobserved.

The parametric bootstrap approach in Björkwall *et al.* (2009) can now be implemented for the separation method using (3.1) and (3.3) in the following way. We draw N_{ij}^* and N_{ij}^{**} from

$$N_{ij}^* \in Po(\hat{n}_{ij}) \quad \text{and} \quad N_{ij}^{**} \in Po(\hat{n}_{ij}) \tag{3.7}$$

B times for all $i, j \in \nabla$ and $i, j \in \Delta$, respectively. We thereby get the B pseudo-triangles ∇N^* and ΔN^{**}. The ultimate number of claims per origin year in the bootstrap world is then given by

$$N_i^{**} = \sum_{j \in \nabla_i} N_{ij}^* + \sum_{j \in \Delta_i} N_{ij}^{**} \tag{3.8}$$

according to (2.1).

Once N_i^{**} is calculated, C_{ij}^* is sampled B times from

$$C_{ij}^* \,|\, N_i^{**} \in \Gamma\left(\frac{N_i^{**}}{\hat{\phi}}, \hat{r}_j \hat{\lambda}_k \hat{\phi}\right), \tag{3.9}$$

for all $i, j \in \nabla$ yielding the B pseudo-triangles ∇C^*. Here $\hat{\lambda}_k$ and $\hat{r}_j$ are obtained from (2.15) and (2.16).

The heuristic estimation process described in Section 2 is then repeated B times for each pair of pseudo-triangles. The claim counts are first forecasted by $\Delta \hat{n}^*$, obtained by the chain-ladder from ∇N^*, in order to estimate the ultimate number of claims per origin year

$$\hat{N}_i^* = \sum_{j \in \nabla_i} N_{ij}^* + \sum_{j \in \Delta_i} \hat{n}_{ij}^* \tag{3.10}$$

according to (2.2). The future payments are then forecasted by estimating $\Delta\hat{m}^*$ according to (2.12)-(2.17). Finally, estimators for the outstanding claims in the bootstrap world are obtained by $\hat{R}_i^* = \sum_{j\in\Delta_i}\hat{m}_{ij}^*$ and $\hat{R}^* = \sum_{\Delta}\hat{m}_{ij}^*$.

In order to generate a random outcome of the true outstanding claims in the bootstrap world, i.e. $R_i^{**} = \sum_{j\in\Delta_i} C_{ij}^{**}$ and $R^{**} = \sum_{\Delta} C_{ij}^{**}$, we sample once again from (3.9) for all $i, j \in \Delta$ to get B triangles ΔC^{**}.

Note that the distribution of C_{ij}^{**} is parameterized based on the estimates of r_j and λ_k which, in the bootstrap world, according to the separation method, are considered as constants instead of stochastic variables. Hence, the C_{ij}^{**} are sampled independently of each other between the rows, while the sampling is independent conditional on N_i within the rows. This implies stochastic independence between the R_i^{**}.

The final step is to calculate the B prediction errors and in Pinheiro *et al.* (2003) this is done by the following equations

$$\mathrm{pe}_i^{**} = \frac{R_i^{**} - \hat{R}_i^*}{\sqrt{\widehat{\mathrm{Var}}(R_i^{**})}} \quad \text{and} \quad \mathrm{pe}^{**} = \frac{R^{**} - \hat{R}^*}{\sqrt{\widehat{\mathrm{Var}}(R^{**})}}. \tag{3.11}$$

The predictive distributions of the outstanding claims R_i and R are then obtained by plotting

$$\tilde{R}_i^{**} = \hat{R}_i + \mathrm{pe}_i^{**}\sqrt{\widehat{\mathrm{Var}}(R_i)} \quad \text{and} \quad \tilde{R}^{**} = \hat{R} + \mathrm{pe}^{**}\sqrt{\widehat{\mathrm{Var}}(R)} \tag{3.12}$$

for each B.

By the conditional independence of C_{ij} for all i and j given N_i (3.3) implies that

$$\begin{aligned}
\mathrm{Var}(R_i) &= E\big(\mathrm{Var}(R_i \mid N_i)\big) + \mathrm{Var}\big(E(R_i \mid N_i)\big) \\
&= E\Big(\sum_{j\in\Delta_i} \phi N_i \big(r_j\lambda_k\big)^2\Big) + \mathrm{Var}\Big(\sum_{j\in\Delta_i} N_i\, r_j\lambda_k\Big) \\
&= \phi E\big(N_i\big)\sum_{j\in\Delta_i}\big(r_j\lambda_k\big)^2 + \mathrm{Var}\big(N_i\big)\Big(\sum_{j\in\Delta_i} r_j\lambda_k\Big)^2 \\
&= \Big(\phi\sum_{j\in\Delta_i}\big(r_j\lambda_k\big)^2 + \Big(\sum_{j\in\Delta_i} r_j\lambda_k\Big)^2\Big)\Big(\sum_{j\in\nabla_i\cup\Delta_i} n_{ij}\Big)
\end{aligned} \tag{3.13}$$

since

$$E(N_i) = \mathrm{Var}(N_i) = \sum_{j\in\nabla_i\cup\Delta_i} n_{ij}\,. \tag{3.14}$$

By plugging in the estimates we find

$$\widehat{\mathrm{Var}}(R_i) = \Big(\hat{\phi}\sum_{j\in\Delta_i}\big(\hat{r}_j\hat{\lambda}_k\big)^2 + \Big(\sum_{j\in\Delta_i}\hat{r}_j\hat{\lambda}_k\Big)^2\Big)\Big(\sum_{j\in\nabla_i\cup\Delta_i}\hat{n}_{ij}\Big) \tag{3.15}$$

and

$$\widehat{\mathrm{Var}}(R) = \sum_i \left(\hat{\phi} \sum_{j \in \Delta_i} \left(\hat{r}_j \hat{\lambda}_k \right)^2 + \left(\sum_{j \in \Delta_i} \hat{r}_j \hat{\lambda}_k \right)^2 \right) \left(\sum_{j \in \nabla_i \cup \Delta_i} \hat{n}_{ij} \right). \tag{3.16}$$

Analogously, the variances appearing in (3.11) are

$$\widehat{\mathrm{Var}}(R_i^{**}) = \left(\hat{\phi}^* \sum_{j \in \Delta_i} \left(\hat{r}_j^* \hat{\lambda}_k^* \right)^2 + \left(\sum_{j \in \Delta_i} \hat{r}_j^* \hat{\lambda}_k^* \right)^2 \right) \left(\sum_{j \in \nabla_i \cup \Delta_i} \hat{n}_{ij}^* \right) \tag{3.17}$$

and

$$\widehat{\mathrm{Var}}(R^{**}) = \sum_i \left(\hat{\phi}^* \sum_{j \in \Delta_i} \left(\hat{r}_j^* \hat{\lambda}_k^* \right)^2 + \left(\sum_{j \in \Delta_i} \hat{r}_j^* \hat{\lambda}_k^* \right)^2 \right) \left(\sum_{j \in \nabla_i \cup \Delta_i} \hat{n}_{ij}^* \right), \tag{3.18}$$

where

$$\hat{\phi}^* = \frac{1}{|\nabla| - q} \sum_{\nabla} \hat{N}_i^* \frac{\left(C_{ij}^* - \hat{N}_i^* \hat{r}_j^* \hat{\lambda}_k^* \right)^2}{\left(\hat{N}_i^* \hat{r}_j^* \hat{\lambda}_k^* \right)^2} \tag{3.19}$$

in accordance with (3.6).

It is remarked in Björkwall *et al.* (2009) that for many bootstrap procedures, resampling of standardized quantities often increases accuracy compared to using unstandardized quantities. Nevertheless, the unstandardized prediction errors

$$\mathrm{pe}_i^{**} = R_i^{**} - \hat{R}_i^* \quad \text{and} \quad \mathrm{pe}^{**} = R^{**} - \hat{R}^* \tag{3.20}$$

are useful, e.g. for the purpose of studying the estimation and the process errors, but also since they are always defined.

The predictive distributions of the outstanding claims R_i and R are then obtained by plotting the B quantities

$$\tilde{R}_i^{**} = \hat{R}_i + \mathrm{pe}_i^{**} \quad \text{and} \quad \tilde{R}^{**} = \hat{R} + \mathrm{pe}^{**}. \tag{3.21}$$

The parametric predictive bootstrap procedure is described in Figure 1 and according to Björkwall *et al.* (2009) we will refer to it as standardized or unstandardized depending on which prediction errors that are used.

3.2. Known claim counts

In Section 2 it was remarked that the separation model is based on the assumption that N_i is considered as known at the moment when the reserving is being

STAGE 1 – REAL WORLD

Substage 1.1 – The triangle of claim counts ∇N

- Forecast the future expected values $\Delta\hat{n}$ and calculate the fitted values $\nabla\hat{n}$ by chain-ladder.
- Calculate the estimated ultimate claim count per origin year $\hat{N}_i$.

Substage 1.2 – The triangle of paid claims ∇C

- Use $\hat{N}_i$ from Substage 1.1 for the purpose of forecasting the future expected values $\Delta\hat{m}$ and calculating the fitted values $\nabla\hat{m}$ by the separation method.
- Calculate $\hat{\phi}$ for the sampling process.
- Estimate the outstanding claims by $\hat{R}_i = \Sigma_{j\in\Delta_i}\hat{m}_{ij}$ and $\hat{R} = \Sigma_{\Delta}\hat{m}_{ij}$.

STAGE 2 – BOOTSTRAP WORLD

Substage 2.1 – The estimated outstanding claims

Substage 2.1.1 – The pseudo-triangle of claim counts ∇N^*

- Sample from (3.7) for $i,j\in\nabla$ to obtain the pseudo-reality in ∇N^*.
- Forecast the future expected values $\Delta\hat{n}^*$ by chain-ladder.
- Calculate the estimated ultimate claim count per origin year $\hat{N}_i^*$.

Substage 2.1.2 – The pseudo-triangle of paid claims ∇C^*

- Sample from (3.7) for $i,j\in\Delta$ to obtain the pseudo-reality in ΔN^{**}.
- Calculate the ultimate claim count per origin year N_i^{**} using ∇N^* from Substage 2.1.1 and ΔN^{**}.
- Sample from (3.9) for $i,j\in\nabla$ to obtain the pseudo-reality in ∇C^* conditionally on ΔN_i^{**}.
- Use $\hat{N}_i^*$ from Substage 2.1.1. for the purpose of forecasting the future expected values $\Delta\hat{m}^*$ by the separation method.
- Estimate the outstanding claims by $\hat{R}_i^* = \Sigma_{j\in\Delta_i}\hat{m}_{ij}^*$ and $\hat{R}^* = \Sigma_{\Delta}\hat{m}_{ij}^*$.

Substage 2.2 – The true outstanding claims

- Sample from (3.9) for $i,j\in\Delta$ to obtain the pseudo-reality in ΔC^{**} conditionally on ΔN_i^{**}.
- Calculate the true outstanding claims $R_i^{**} = \Sigma_{j\in\Delta_i}C_{ij}^{**}$ and $R^{**} = \Sigma_{\Delta}C_{ij}^{**}$.
- Store either the standardized prediction errors in (3.11) or the unstandardized ones in (3.20).
- Return to the beginning of the bootstrap loop in Stage 2 and repeat B times.

STAGE 3 – ANALYSIS OF THE SIMULATIONS

- Obtain the predictive distribution of R_i and R, the true outstanding claims in the real world, by plotting the B values in either (3.12) or (3.21).

FIGURE 1: The procedure of the parametric predictive bootstrap for the separation method.

done. This can often be a reasonable assumption since the numbers of claims are usually finalized quite early even for long-tailed business. In Section 3.1 N_i was a random variable; in order to get a view of how much uncertainty N_i contributes to the predictive distribution of the claims reserve we now consider the special case when N_i is treated as deterministic, in contrast to (3.1). Consequently, $\hat{N}_i \equiv N_i$ in all equations above.

Assumption (3.3) can still be used and $\hat{\phi}$ is estimated as in (3.6), but the sampling process changes. We do not have to generate pseudo-triangles of claim counts in the bootstrap world, i.e. ∇N^* and ΔN^{**}, since N_i is considered as known. Thus, we just draw C_{ij}^* from

$$C_{ij}^* \in \Gamma\left(\frac{N_i}{\hat{\phi}}, \hat{r}_j \hat{\lambda}_k \hat{\phi}\right) \tag{3.22}$$

B times for all $i, j \in \nabla$ yielding ∇C^*. The estimation process of the separation method is then repeated for each ∇C^* using N_i as the exposure in the bootstrap world as well. Finally, we sample once again B times from (3.22) for all $i, j \in \Delta$ to get ΔC^{**}.

The prediction errors and the predictive distributions are as earlier obtained by (3.11) and (3.12), respectively, but since $\mathrm{Var}(N_i) = 0$, we obtain the estimators

$$\widehat{\mathrm{Var}}(R_i) = \hat{\phi} N_i \sum_{\Delta_i} \left(\hat{r}_j \hat{\lambda}_k\right)^2 \tag{3.23}$$

and

$$\widehat{\mathrm{Var}}(R) = \sum_i \hat{\phi} N_i \sum_{\Delta_i} \left(\hat{r}_j \hat{\lambda}_k\right)^2 \tag{3.24}$$

instead of (3.15) and (3.16).

Analogously, the estimators appearing in (3.11) are

$$\widehat{\mathrm{Var}}(R_i^{**}) = \hat{\phi}^* N_i \sum_{\Delta_i} \left(\hat{r}_j^* \hat{\lambda}_k^*\right)^2 \tag{3.25}$$

and

$$\widehat{\mathrm{Var}}(R^{**}) = \sum_i \hat{\phi}^* N_i \sum_{\Delta_i} \left(\hat{r}_j^* \hat{\lambda}_k^*\right)^2, \tag{3.26}$$

where $\hat{\phi}^*$ is estimated by (3.19) with $\hat{N}_i^*$ replaced by N_i.

The unstandardized prediction errors in (3.20) can of course be used as well. The predictive distributions are then obtained by (3.21).

This simplified approach is summarized in Figure 2.

STAGE 1 – REAL WORLD

Substage 1.1 – The triangle of claim counts ∇N

- Forecast the future expected values $\Delta\hat{n}$ by chain-ladder.
- Calculate the estimated ultimate claim count per origin year $\hat{N}_i$.

Substage 1.2 – The triangle of paid claims ∇C

- Use $\hat{N}_i$ from Substage 1.1 for the purpose of forecasting the future expected values $\Delta\hat{m}$ and calculating the fitted values $\nabla\hat{m}$ by the separation method.
- Calculate $\hat{\phi}$ for the sampling process.
- Calculate the outstanding claims $\hat{R}_i = \Sigma_{j\in\Delta_i}\hat{m}_{ij}$ and $\hat{R} = \Sigma_{\Delta}\hat{m}_{ij}$.

STAGE 2 – BOOTSTRAP WORLD

Substage 2.1 – The estimated outstanding claims

- Sample from (3.22) for $i,j \in \nabla$ to obtain the pseudo-reality in ∇C^*.
- Use $\hat{N}_i$ for the purpose of forecasting the future expected values $\Delta\hat{m}^*$ by the separation method.
- Calculate the estimated outstanding claims $\hat{R}_i^* = \Sigma_{j\in\Delta_i}\hat{m}_{ij}^*$ and $\hat{R}^* = \Sigma_{\Delta}\hat{m}_{ij}^*$.

Substage 2.2 – The true outstanding claims

- Sample from (3.22) for $i,j \in \Delta$ to obtain the pseudo-reality in ΔC^{**}.
- Calculate the true outstanding claims $R_i^{**} = \Sigma_{j\in\Delta_i}C_{ij}^{**}$ and $R^{**} = \Sigma_{\Delta}C_{ij}^{**}$.
- Store either the standardized prediction errors in (3.11) or the unstandardized ones in (3.20).
- Return to the beginning of the bootstrap loop in Stage 2 and repeat B times.

STAGE 3 – ANALYSIS OF THE SIMULATIONS

- Obtain the predictive distribution of R_i and R, the true outstanding claims in the real world, by plotting the B values in either (3.12) or (3.21).

FIGURE 2: The procedure of the simplified parametric predictive bootstrap for the separation method.

4. NUMERICAL STUDY

The purpose of the numerical study is to illustrate the parametric bootstrap procedure for the separation method and to compare it to the approach for the chain-ladder described in Björkwall *et al.* (2009). From now on $B = 10\,000$ simulations are used for each prediction. The upper 95 percent limits are studied and the coefficients of variation, i.e. $\sqrt{\mathrm{Var}(\tilde{R}_i^{**})}/\hat{R}_i$ and $\sqrt{\mathrm{Var}(\tilde{R}^{**})}/\hat{R}$, are presented as well.

We use the well-known data from Taylor & Ashe (1983), who also provide observations of number of claims. The triangles of paid claims ∇C and claim counts ∇N are presented in Table 4.1 and Table 4.2, respectively.

TABLE 4.1

OBSERVATIONS OF PAID CLAIMS ∇C FROM TAYLOR & ASHE (1983)

	0	1	2	3	4	5	6	7	8	9
0	357 848	766 940	610 542	482 940	527 326	574 398	146 342	139 950	227 229	67 948
1	352 118	884 021	933 894	1 183 289	445 745	320 996	527 804	266 172	425 046	
2	290 507	1 001 799	926 219	1 016 654	750 816	146 923	495 992	280 405		
3	310 608	1 108 250	776 189	1 562 400	272 482	352 053	206 286			
4	443 160	693 190	991 983	769 488	504 851	470 639				
5	396 132	937 085	847 498	805 037	705 960					
6	440 832	847 631	1 131 398	1 063 269						
7	359 480	1 061 648	1 443 370							
8	376 686	986 608								
9	344 014									

TABLE 4.2

OBSERVATIONS OF CLAIM COUNTS ∇N FROM TAYLOR & ASHE (1983)

	0	1	2	3	4	5	6	7	8	9
0	40	124	157	93	141	22	14	10	3	2
1	37	186	130	239	61	26	23	6	6	
2	35	158	243	153	48	26	14	5		
3	41	155	218	100	67	17	6			
4	30	187	166	120	55	13				
5	33	121	204	87	37					
6	32	115	146	103						
7	43	111	83							
8	17	92								
9	22									

4.1. The estimated claims reserve

The assumption of the future inflation rate has great impact on the claims reserve estimated by the separation method. The future inflation rate can of course be modeled by more refined approaches, but this is beyond the scope of this paper and, hence, we just consider two simple models. The first one is to use the mean rate observed so far, $\hat{K}_{\text{mean}} = 11{,}01\%$, in (2.18) and the second one is to estimate K in $\lambda_k = (1 + K)^k \lambda_0$ by loglinear regression and then use this estimator $\hat{K}_{\text{reg}} = 9{,}33\%$ in (2.18). It is not appropriate to adopt the loglinear regression curve as a whole for the future inflation rate, since we then end up with a discontinuity point in λ_t. Thus, we only keep the term corresponding to the slope

TABLE 4.3

THE ESTIMATED CLAIMS RESERVES UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WITH DIFFERENT INFLATION ASSUMPTIONS $\hat{K}_{\text{mean}}$ AND $\hat{K}_{\text{reg}}$

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Chain-ladder
1	89 163	87 817	94 634
2	506 151	497 097	469 511
3	794 132	773 286	709 638
4	1 288 308	1 246 830	984 889
5	1 722 883	1 657 265	1 419 459
6	2 448 039	2 344 265	2 177 641
7	3 269 931	3 128 918	3 920 301
8	4 314 184	4 112 206	4 278 972
9	6 043 441	5 726 918	4 625 811
Total	20 476 232	19 574 602	18 680 856

as the future inflation rate. Table 4.3 presents the estimated reserves for the separation method under the two assumptions as well as for the chain-ladder.

4.2. Predictive bootstrap results for the chain-ladder

In order to compare the separation method to the chain-ladder we summarize the results of the parametric predictive bootstrap procedures described in Björkwall *et al.* (2009), where data is bootstrapped according to the plug-in-principle under the assumption of a gamma distribution; see the reference for details.

TABLE 4.4

THE 95 PERCENTILES OF THE PARAMETRIC PREDICTIVE BOOTSTRAP PROCEDURES DESCRIBED IN BJÖRKWALL *ET AL.* (2009) FOR THE CHAIN-LADDER. WE WORK UNDER THE ASSUMPTION OF A GAMMA DISTRIBUTION AND THE PROCEDURE IS EITHER STANDARDIZED OR UNSTANDARDIZED.

Year i	Standardized Gamma	Unstandardized Gamma
1	219 178	168 756
2	861 781	756 634
3	1 169 041	1 062 783
4	1 519 540	1 409 034
5	2 127 947	1 975 222
6	3 358 037	3 038 732
7	6 253 164	5 562 133
8	7 386 412	6 284 020
9	9 247 043	7 148 120
Total	23 991 467	23 123 593

TABLE 4.5

THE COEFFICIENTS OF VARIATION OF THE SIMULATIONS (IN %) OF THE PARAMETRIC PREDICTIVE BOOTSTRAP PROCEDURES DESCRIBED IN BJÖRKWALL *ET AL.* (2009) FOR THE CHAIN-LADDER. WE WORK UNDER THE ASSUMPTION OF A GAMMA DISTRIBUTION AND THE PROCEDURE IS EITHER STANDARDIZED OR UNSTANDARDIZED

Year i	Standardized Gamma	Unstandardized Gamma
1	65	50
2	41	38
3	32	31
4	28	27
5	26	25
6	27	25
7	29	27
8	35	32
9	47	38
Total	15	16

Tables 4.4 - 4.5 show the results for the standardized as well as the unstandardized approach.

Note that the 95 percentiles of the total reserve in Table 4.4, according to the standardized and the unstandardized approaches, respectively, are quite close even though the percentiles of some of the accident years differ radically. The difference is largest for the latest accident years due to their sensitivity to occasional extreme observations. Hence, $\widehat{\mathrm{Var}}(R_i^{**})$ in (3.11) becomes large which implies that the predictive distribution of the standardized prediction error is being compressed compared to the unstandardized case, see Section 4.4 and Figures 3(c)-(d) for further details. The difference between the two approaches are not that large for the total chain-ladder reserve, since it is more robust to extreme observations compared to an individual accident year.

4.3. The standardized predictive bootstrap for the separation method

The results for the bootstrap procedure described in Section 3.1, when the standardized prediction errors are used, are presented in Table 4.6 for four different assumptions of the future inflation rate. Two of these are mean inflation rates observed so far, either treated as a constant, $\hat{K}_{\text{mean}}$, or as stochastic in the bootstrap world, $\hat{K}^*_{\text{mean}}$. The other two correspond to the inflation rate estimated by loglinear regression, either treated as a constant, $\hat{K}_{\text{reg}}$, or as stochastic in the bootstrap world, $\hat{K}^*_{\text{reg}}$. According to the plug-in-principle the inflation rate should be treated as stochastic, i.e. recomputed from $\{\hat{\lambda}^*_k\}$ for each resample, but the constant alternatives are shown as well for comparison. Table 4.7 contains the coefficients of variation. Tables 4.6 - 4.7 also include the results obtained by the chain-ladder for comparison.

Since the two inflation assumptions $\hat{K}_{\text{mean}} = 11{,}01\%$ and $\hat{K}_{\text{reg}} = 9{,}33\%$ are quite close to each other it is hard to find any reliable differences in the 95 percent limits in Table 4.6 due to the variation inherent in estimating the tails of a distribution. However, the coefficients of variation in Table 4.7 are higher when the inflation is treated as stochastic in the bootstrap world, in particular for the grand total. As expected the coefficients of variation of the latest origin year are lower for the separation method than for the chain-ladder, since the extreme sensitivity to outliers for the chain-ladder in the south corner of the upper triangle is removed for the separation method. Less expected is that the separation method has lower coefficients of variation for years 1-4.

TABLE 4.6

THE 95 PERCENTILES OF THE STANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WITH FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	201 184	199 734	205 997	197 700	219 178
2	882 300	859 549	880 776	883 330	861 781
3	1 253 445	1 214 151	1 221 817	1 219 214	1 169 041
4	1 908 980	1 863 782	1 853 224	1 880 126	1 519 540
5	2 513 476	2 443 947	2 405 583	2 471 176	2 127 947
6	3 526 976	3 481 354	3 384 454	3 488 347	3 358 037
7	4 893 910	4 770 103	4 675 530	4 881 698	6 253 164
8	6 540 182	6 371 191	6 223 934	6 481 017	7 386 412
9	9 852 469	9 598 085	9 426 785	9 473 484	9 247 043
Total	27 442 696	27 661 199	26 229 943	27 267 855	23 991 467

TABLE 4.7

THE COEFFICIENTS OF VARIATION OF THE SIMULATIONS (IN %) OF THE STANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WITH FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	61	60	63	62	65
2	38	37	38	39	41
3	29	30	30	30	32
4	25	27	26	26	28
5	24	27	24	26	26
6	23	27	23	26	27
7	26	30	26	29	29
8	27	32	27	30	35
9	32	37	32	35	47
Total	18	24	18	21	15

4.4. The unstandardized predictive bootstrap for the separation method

In order to study the estimation and the process error we also investigate the procedure described in Section 3.1 when the unstandardized prediction errors are used. The results are shown in Tables 4.8 - 4.9.

As remarked in Björkwall *et al.* (2009) the percentiles of the unstandardized predictive bootstrap tend to be lower than for the standardized one. This was explained by the left skewness of the predictive distribution of the unstandardized bootstrap compared to the distribution obtained by the standardized bootstrap. According to Figure 3 this seems to hold for the separation method too.

TABLE 4.8

THE 95 PERCENTILES OF THE UNSTANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WITH FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	158 866	154 338	156 499	155 430	168 756
2	803 966	792 594	797 648	799 394	756 634
3	1 180 997	1 140 365	1 157 541	1 157 309	1 062 783
4	1 825 048	1 779 729	1 777 252	1 783 739	1 409 034
5	2 413 236	2 354 841	2 317 790	2 337 771	1 975 222
6	3 389 043	3 334 188	3 272 652	3 287 929	3 038 732
7	4 666 419	4 537 282	4 491 341	4 520 117	5 562 133
8	6 180 526	6 068 741	5 945 143	5 992 772	6 284 020
9	9 024 898	8 778 894	8 542 080	8 670 774	7 148 120
Total	26 091 962	26 127 705	25 020 103	25 589 970	23 123 593

TABLE 4.9

THE COEFFICIENTS OF VARIATION OF THE SIMULATIONS (IN %) OF THE UNSTANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WITH FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	48	50	49	48	50
2	36	38	36	36	38
3	29	33	29	30	31
4	25	32	25	27	27
5	24	34	24	26	25
6	23	37	24	26	25
7	26	40	26	28	27
8	26	43	27	30	32
9	32	53	32	36	38
Total	18	36	18	22	16

Figures 3(c)-(d) show the predictive distributions of the total claims reserve under the assumption of a stochastic future inflation rate corresponding to the mean inflation rate observed so far. The predictive distribution obtained by the unstandardized bootstrap in (c) is skewed to the left compared to the one obtained by the standardized bootstrap in (d), which is slightly skewed to the right. This follows since the process component in Figure 3(a) has smaller variability than the estimation component in Figure 3(b), and the latter is skewed to the right. The left skewness is to a large extent removed for the standardized prediction errors (3.11), because of the denominator, but not for the unstandardized prediction errors (3.20).

Recomputing the future inflation rate from $\{\hat{\lambda}_k^*\}$ for each resample in the bootstrap world yields some rates which are unreasonably high when we use $\hat{K}_{\text{mean}}^*$. These rates stretch the tail of the estimation error component to the right. Consequently, the predictive distribution of the outstanding claims is more stretched to the left for the stochastic future inflation rate than for the constant. This explains why most of the percentiles in Tables 4.6 and 4.8 are

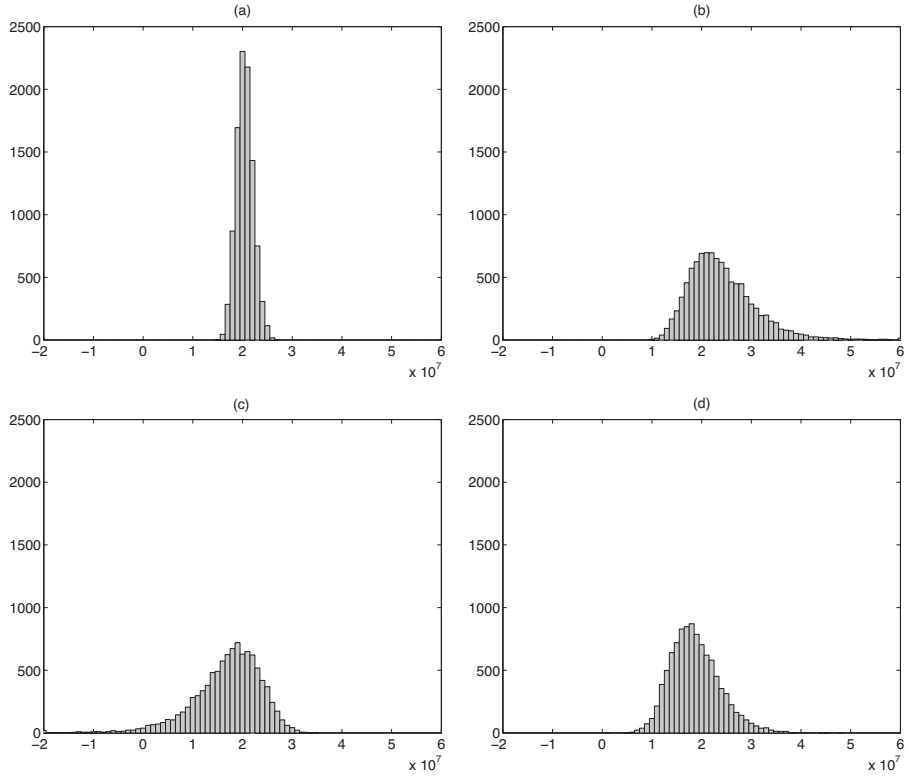

FIGURE 3: Density charts of R^{**} (a), $\hat{R}^*$ (b) and $\tilde{R}^{**}$ for the unstandardized (c) and standardized (d) predictive bootstrap procedure under the assumption of a stochastic future inflation rate corresponding to the mean inflation rate observed so far.

lower for $\hat{K}^*_{\text{mean}}$ than $\hat{K}_{\text{mean}}$. The assumption of $\hat{K}^*_{\text{reg}}$ seem to be more stable, which is also confirmed by the coefficients of variation in Table 4.7 and Table 4.9.

In Table 4.9 we can see that the coefficients of variation for the chain-ladder are generally higher than those of the separation method with fixed inflation for individual accident years, yet less in total. Moreover, from Table 4.8 it can be concluded that the diversification effect on the total is larger for the chain-ladder than for the separation method. Hence, the correlations between the rows must be lower for the chain-ladder than for the separation method and this may occur because of the λ_k which diagonally affects all rows.

TABLE 4.10

THE 95 PERCENTILES OF THE SIMPLIFIED STANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WHEN N_i IS CONSIDERED AS KNOWN. WE WORK UNDER FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	203 666	194 659	196 435	197 139	219 178
2	897 770	875 334	878 776	894 381	861 781
3	1 243 302	1 195 689	1 226 422	1 228 154	1 169 041
4	1 918 643	1 858 748	1 875 831	1 879 533	1 519 540
5	2 525 290	2 451 200	2 433 665	2 468 674	2 127 947
6	3 577 632	3 481 789	3 413 650	3 517 184	3 358 037
7	4 871 638	4 819 584	4 720 788	4 830 392	6 253 164
8	6 507 485	6 421 424	6 216 127	6 423 967	7 386 412
9	9 023 231	8 859 782	8 505 990	9 057 342	9 247 043
Total	27 417 470	27 783 761	26 290 837	27 504 922	23 991 467

TABLE 4.11

THE COEFFICIENTS OF VARIATION OF THE SIMULATIONS (IN %) OF THE SIMPLIFIED STANDARDIZED PREDICTIVE BOOTSTRAP PROCEDURE UNDER THE CHAIN-LADDER, COMPARED TO THE SEPARATION METHOD WHEN N_i IS CONSIDERED AS KNOWN. WE WORK UNDER FOUR DIFFERENT INFLATION ASSUMPTIONS

Year i	Inflation $\hat{K}_{\text{mean}}$	Inflation $\hat{K}^*_{\text{mean}}$	Inflation $\hat{K}_{\text{reg}}$	Inflation $\hat{K}^*_{\text{reg}}$	Chain Ladder Gamma
1	62	57	60	62	65
2	39	39	39	39	41
3	29	30	30	30	32
4	26	27	26	26	28
5	24	27	24	26	26
6	24	27	23	26	27
7	26	31	26	28	29
8	26	32	26	29	35
9	25	32	25	29	47
Total	17	24	17	21	15

4.5. Known claim counts

In Tables 4.10 - 4.11 we present the results of the simplified approach in Section 3.2 where we treat N_i as known. Again it is hard to conclude whether or not there are any reliable differences compared to Tables 4.6 - 4.7 except for the last origin year, i.e. where we predict the ultimate number of claims based on one single observation. However, the notably small difference is consistent with the separation method assumption that the numbers of claims usually are finalized early enough to be considered as known. This is interesting, since Table 4.2 reveals that the data used here is actually an example when claim numbers are not finalized very fast. Of course, the situation might be different in another example.

5. Discussion

The separation method was, like the chain-ladder, originally formulated as a deterministic method. The chain-ladder was given a stochastic interpretation by Mack (1993) who analytically calculated the MSEP for the claims reserve using a second-moment assumption. England & Verrall (1999), England (2002) and Pinheiro *et al.* (2003) introduced bootstrapping in order to provide a full predictive distribution of the claims reserve using a GLM framework and, recently, Björkwall *et al.* (2009) suggested an alternative parametric bootstrap approach for which the GLM assumption has been relaxed. However, in contrast to the development of the chain-ladder method this paper directly gives the separation method a parametric formulation in accordance with Björkwall *et al.* (2009).

Indeed, the alternative to resampling would be to analytically calculate the MSEP for the separation method reserve estimate. This approach is computationally more feasible, but requires an explicit expression of the MSEP as well as further distributional assumptions in order to obtain a full predictive distribution. For the chain-ladder method, Mack (1993) derived the MSEP within an autoregressive formulation of the claims development, whereas Renshaw (1994) and England & Verrall (1999) derived a first order Taylor approximation of the chain-ladder MSEP within a GLM framework. A corresponding Taylor approximation for the separation method would be

$$MSEP(R) = \mathrm{Var}(R) + \sum_{i_1, j_1 \in \Delta} \sum_{i_2, j_2 \in \Delta} \mathrm{Cov}\big(\hat{\mu}_{i_1, j_1}, \hat{\mu}_{i_2, j_2}\big), \tag{5.1}$$

where $\mathrm{Var}(R)$ is deduced as in Section 3.1 and $\hat{\mu}_{ij}$ is an estimate of $\mu_{ij} = E(N_i)\, r_j\, \lambda_k$. However, calculation of the covariance terms of (5.1) requires derivation of the asymptotic covariance matrix of estimates of all r_j, λ_k and the parameters involved in $E(N_i)$. This is quite a challenging task, since we cannot use GLM theory directly, as will now be explained.

In the parametric bootstrap approach for the separation method described in Section 3.1 we have chosen to use Poisson distributed claim counts and gamma distributed claim amounts conditional on N_i in (3.1) and (3.2), respectively. Unconditionally, this implies compound Poisson distributed claim amounts. Jørgensen & de Souza (1994) and Smyth & Jørgensen (2002) have shown that the compound Poisson distribution belongs to the Tweedie class within the exponential dispersion family, with variance function $V(\mu) = \mu^p$ for some $1 < p < 2$. However, we cannot directly use GLM theory for inference, since all C_{ij} within the same row of the development triangle share the same N_i and thus are dependent.

Other distributions of $C_{ij}|N_i$ than gamma might be considered, but the gamma class is fairly flexible, involving both scale and shape parameters, and is guaranteed to be positive. In addition, the convolution property (3.2)-(3.3) of the gamma class of distribution makes it particularly appealing. Furthermore, a more general class of distributions would be obtained by choosing N_i from a mixed Poisson distribution. This is an interesting possibility. Over-dispersed Poisson distributions (ODP) are often used in a GLM regression context where only the variance function, not the entire distribution, is required. However, when the entire distribution is needed, as in (3.1), the ODP is not an adequate model for claim counts since its support is not on the non-negative numbers; hence, we believe that it would be more relevant to adopt a mixed Poisson distribution if one wants to introduce a dispersion parameter for N_i.

An alternative non-parametric bootstrap procedure could be defined by removing (3.3) and using only the first and second order assumptions (3.4)-(3.5). This yields, conditional on N_i, standardized residuals

$$r_{ij} = \frac{C_{ij} - \widehat{E}(C_{ij}|N_i)}{\sqrt{\widehat{\mathrm{Var}}(C_{ij}|N_i)}} \tag{5.2}$$

to resample from. This resampling could then be combined with the current resampling of claim counts N_{ij}, which in fact does not require the Poisson assumption, but only assumptions on the first two moments of all N_{ij}, see Björkwall *et al.* (2009) and references therein. For the non-parametric bootstrap approach the second moment could be the variance corresponding to an ODP. It would be interesting to compare the parametric and non-parametric resampling procedures in a separate paper.

The issue of whether standardized or unstandardized prediction errors would be preferable in this context has not yet been discussed in the literature as far as we are aware of. For instance, Pinheiro *et al.* (2003) use standardized prediction errors according to the general bootstrap procedure in Davison & Hinkley (1997), while England (2002) and Li (2006) use unstandardized ones. Björkwall *et al.* (2009) attempt to approach this issue by noting that resampling of standardized quantities often increases accuracy compared to using unstandardized quantities, and, hence, the standardized case is in theory preferable

even though the unstandardized case might be more useful in practice since it is easier to implement. Note that Björkwall *et al.* (2009) suggest a double bootstrap approach for the standardization, since the standard deviation used for the prediction errors in Pinheiro *et al.* (2003) is an approximation.

A disadvantage of the standardized prediction errors (3.11) for R_i and R is that we cannot aggregate the R_is to arrive at R, since each bootstrapped quantity should be standardized individually. Hence, we cannot study the diversification effects on the grand total R for the standardized approach, while this is not a problem for the unstandardized case where we use the prediction errors in (3.20).

Note that it is possible to form the standardized as well as the unstandardized prediction errors for each single cell in Δ if needed as a complement to (3.11) and (3.20), respectively. However, again we cannot add the cells within a row Δ_i up to obtain R_i for the standardized bootstrap, since R_i should be standardized individually. Despite the lack of this feature it is still possible to include operations regarding the payment patterns, e.g. discounting of the cash flows, since they can be modeled within the bootstrap procedure, see Björkwall *et al.* (2009) for details.

Acknowledgements

The authors are grateful to Björn Johansson, Länsförsäkringar Alliance, for his interpretation (2.8)-(2.9) of individually paid claims, which inspired our suggested resampling method.

References

Björkwall, S., Hössjer, O. and Ohlsson, E. (2009) Non-parametric and Parametric Bootstrap Techniques for Age-to-Age Development Factor Methods in Stochastic Claims Reserving. *Scandinavian Actuarial Journal*, **4**, 306-331.

Buchwalder, M., Bühlmann, H., Merz, M. and Wüthrich, M. (2006) The Mean Square Error of Prediction in the Chain Ladder Reserving Method (Mack and Murphy revisited). *Astin Bulletin*, **36(2)**, 521-542.

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

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

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

Jørgensen, B. and De Souza, M.C.P. (1994) Fitting Tweedie's Compound Poisson Model to Insurance Claims Data. *Scandinavian Actuarial Journal*, **1**, 69-93.

Li, J. (2006) Comparison of Stochastic Reserving Methods. *Australian Actuarial Journal*, **12(4)**, 489-569.

Liu, H. and Verrall, R. (2008) *Bootstrap Estimation of the Predictive Distributions of Reserves Using Paid and Incurred Claims.* ASTIN Colloquium, Manchester 2008. Available online at: http://www.actuaries.org/ASTIN/Colloquia/Manchester/Papers_EN.cfm

Mack, T. (1993) Distribution Free Calculation of the Standard Error of Chain Ladder Reserve Estimates. *Astin Bulletin*, **23(2)**, 213-225.

MACK, T. (2008) The Prediction Error of Bornhuetter/Ferguson. *Astin Bulletin*, **38(1)**, 87-103.
MACK, T., QUARG, G. and BRAUN, C. (2006) The Mean Square Error of Prediction in the Chain Ladder Reserving Method – a Comment. *Astin Bulletin*, **36(2)**, 543-552.
MCCULLAGH, P. and NELDER, J.A. (1989) *Generalized Linear Models*. London: Chapman and Hall.
PINHEIRO, P.J.R., ANDRADE E SILVA, J.M. and CENTENO, M.D.L (2003) Bootstrap Methodology in Claim Reserving. *The Journal of Risk and Insurance*, **4**, 701-714.
QUARG, G. and MACK, T. (2004) Munich Chain Ladder. *Blätter der DGVFM*, **26**, 597-630.
RENSHAW, A.E. (1994) Modelling the Claims Process in the Presence of Covariates. *Astin Bulletin*, **24**, 265-286.
SMYTH, G.K. and JØRGENSEN, B. (2002) Fitting Tweedie's Compound Poisson Model to Insurance Claims Data: Dispersion Modelling. *Astin Bulletin*, **32**, 143-157.
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. and ASHE, F.R. (1983) Second Moments of Estimates of Outstanding Claims. *Journal of Econometrics*, **23**, 37-61.
VERBEEK, H.G. (1972) An Approach to the Analysis of Claims Experience in Motor Liability Excess of Loss Reassurance. *Astin Bulletin*, **6**, 195-202.

SUSANNA BJÖRKWALL
Mathematical Statistics, c/o Hössjer
SE-106 91 Stockholm University
E-Mail: sbj@student.su.se