View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by RERO DOC Digital Library

CREDIBILITY FOR THE CHAIN LADDER RESERVING METHOD

BY

ALOIS GISLER AND MARIO V. WÜTHRICH

ABSTRACT

We consider the chain ladder reserving method in a Bayesian set up, which allows for combining the information from a specific claims development triangle with the information from a collective. That is, for instance, to consider simultaneously own company specific data and industry-wide data to estimate the own company's claims reserves. We derive Bayesian estimators and credibility estimators within this Bayesian framework. We show that the credibility estimators are exact Bayesian in the case of the exponential dispersion family with its natural conjugate priors. Finally, we make the link to the classical chain ladder method and we show that using non-informative priors we arrive at the classical chain ladder forecasts. However, the estimates for the mean square error of prediction differ in our Bayesian set up from the ones found in the literature. Hence, the paper also throws a new light upon the estimator of the mean square error of prediction of the classical chain ladder forecasts and suggests a new estimator in the chain ladder method.

KEYWORDS

Chain Ladder, Bayes Statistics, Credibility Theory, Exponential Dispersion Family, Mean Square Error of Prediction.

1. INTRODUCTION

Claims reserving is one of the basic actuarial tasks in the insurance industry. Based on observed claims development figures (complete or incomplete development triangles or trapezoids) actuaries have to predict the ultimate claim amount for different lines of business as well as for the whole insurance portfolio. They are often confronted with the problem that the observed development figures within a given loss development triangle heavily fluctuate due to random fluctuations and a scarce data base. This makes it difficult to make a reliable forecast for the ultimate claim. In such situations, actuaries often rely on industry-wide development patterns rather than on the observed company data. The question then arises, when and to what extent should one rely on the

Astin Bulletin 38(2), 565-600. doi: 10.2143/AST.38.2.2033354 © 2008 by Astin Bulletin. All rights reserved.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

industry-wide data. A similar question arises when considering different lines of business. If the data of a line of business is too scarce, one typically considers the development pattern of other similar lines of business. Here, again, the question occurs how much should one rely on the claims experience of other similar lines of business and how much weight should one give to the observations of the individual line under consideration.

The mathematical tool to answer such kind of questions is credibility theory. The point is that besides the specific claims development triangle there are also other sources of information available (so-called collective information such as the development pattern of industry-wide data, the development figures of other "similar" lines of business or expert opinion), which may tell something about the future development of the claims of the specific claims development triangle in question. Credibility theory allows for modelling such situations and gives an answer to the question of how to combine specific and collective claims information to get a best estimate of the specific ultimate claim amount.

In claims reserving, various models and methods are found in the literature, for an overview we refer to England and Verrall [5] and Wüthrich and Merz [21]. In the following we concentrate on the chain ladder reserving method which is still one of the best known and most popular method in the insurance practice. However, the basic idea of this paper, namely to consider a Bayesian set up and to use credibility techniques for estimating the ultimate claim amount can be transformed to any other claims reserving method.

2. Classical Chain Ladder

Assume that $C_{i,j}$ denotes the total cumulative claim of accident year $i \in \{0, \dots, I\}$ at the end of development period $j \in \{0, \dots, J\}$. Without loss of generality, we assume that the development period is one year. Usually, $C_{i,j}$ denotes either cumulative claims payments or claims incurred, but it could also be another quantity like, for instance, the number of reported claims. We assume that the claims development ends at development year J and that therefore $C_{i,J}$ is the total ultimate claim amount of accident year i. Throughout this paper, $C_{i,j}$ is referred to as claim of accident year i at the end of development year j and $C_{i,J}$ as the ultimate claim. The exact definition of "claim" depends on the situation and on the claims data considered (cumulative payments or claims incurred). At time I we have observations (upper left trapezoid)

$$\mathcal{D}_I = \{C_{i,j} : 0 \le i \le I,\ 0 \le j \le J,\ i + j \le I\}, \tag{2.1}$$

and the random variables $C_{i,j}$ need to be predicted for $i + j > I$. In particular, we want to predict for each accident year i the ultimate claim $C_{i,J}$ and the outstanding loss liabilities (if $C_{i,j}$ refers to payments)

$$R_i = C_{i,J} - C_{i,I-i}. \tag{2.2}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

We define for $j = 0, 1, \ldots, J$

$$\mathbf{C}_j = (C_{0,j}, C_{1,j}, \ldots, C_{I-j,j})', \tag{2.3}$$

the column vectors of the observed trapezoid $\mathcal{D}_I$, and for $k \leq I - j$

$$S_j^{[k]} = \sum_{i=0}^{k} C_{i,j}. \tag{2.4}$$

The chain ladder method was originally understood as a deterministic algorithm for setting claims reserves without having an underlying stochastic model in mind. The basic assumption behind the chain ladder method is that successive column vectors $\{C_{i,j} : i = 0, 1, \ldots, I\}$ for $j = 0, \ldots, J$ are, up to random fluctuations, proportional to each other, i.e.

$$C_{i,j+1} \simeq f_j \, C_{i,j} \tag{2.5}$$

for appropriate constants $f_j > 0$. These factors f_j are called chain ladder factors, development factors or age-to-age factors. The chain ladder algorithm is such that at time I the random variables $C_{i,k}$ for $k > I - i$ are predicted by the chain ladder forecasts

$$C_{i,k}^{CL} = C_{i,I-i} \prod_{j=I-i}^{k-1} \hat{f}_j, \tag{2.6}$$

where

$$\hat{f}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}. \tag{2.7}$$

Thus the ultimate claim $C_{i,J}$ is predicted by $C_{i,J}^{CL}$ and the chain ladder reserve of accident year i at time I is

$$R_i^{CL} = C_{i,J}^{CL} - C_{i,I-i}. \tag{2.8}$$

Remark:

- To be strict, (2.8) is the chain ladder reserve for a cumulative payments triangle or trapezoid. For incurred claims, the chain ladder reserve is

$$\begin{aligned} R_i^{CL} &= C_{i,J}^{CL} - C_{i,I-i}^{paid} \\ &= (C_{i,J}^{CL} - C_{i,I-i}) + (C_{i,I-i} - C_{i,I-i}^{paid}), \end{aligned} \tag{2.9}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

i.e. we have to add the difference between the incurred claims $C_{i,I-i}$ and the cumulative payments $C_{i,I-i}^{paid}$. Both are known at time I, i.e. the difference between (2.9) and (2.8) is a known constant, which has no impact on the reserving problem and the associated uncertainty. Without loss of generality we therefore just consider the reserves as defined in (2.8).

It is the merit of Mack [10] to have formulated the stochastic model underlying the chain ladder method. Mack's model relies on the following model assumptions:

Model Assumptions 2.1. (Mack's chain ladder model)

M1 Random variables $C_{i,j}$ belonging to different accident years $i \in \{0, 1, \ldots, I\}$ are independent.

M2 There exist constants $f_j > 0$ and $\sigma_j^2 > 0$, such that for all $i \in \{0, 1, \ldots, I\}$ and for all $j = 0, 1, \ldots, J-1$ we have

$$E[C_{i,j+1} \mid C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = f_j\, C_{i,j}, \tag{2.10}$$

$$\mathrm{Var}[C_{i,j+1} \mid C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = \sigma_j^2\, C_{i,j}. \tag{2.11}$$

Note that Mack's model is a distribution-free model making only assumptions on the conditional first and second moments.

The advantage of an underlying stochastic model is that it does not only yield a point estimate for the ultimate claim but that it also allows for estimating the standard error of the chain ladder prediction. A formula for estimating the mean square error of prediction was derived in Mack [10]. This formula was the subject of discussions and of further investigations for example in Barnett and Zehnwirth [1], Buchwalder et al. [2], Mack et al. [13], Gisler [6] and Venter [17].

In the literature, one finds also other stochastic models leading to the chain ladder forecasts given by (2.6) and (2.7), in particular the (overdispersed) Poisson model with maximum likelihood estimators leads also to the chain ladder reserves. More details on these models and a comparison of other models related to the chain ladder method can be found in Mack [9], England and Verrall [5], Mack and Venter [12], Hess and Schmidt [7], and Mack [11]. However, in our opinion, Mack's model reflects best the very idea behind chain ladder method.

In this paper we introduce a *Bayesian chain ladder model*, which is *the Bayesian counterpart to Mack's model*. We will see later that under certain conditions and by using non-informative priors, the chain ladder forecasts in the Bayesian model are the same as in the classical chain ladder model. However, the estimators of the mean square error of prediction are different to the ones given in Mack [10]. Moreover, we would like to remark that the Bayesian model considered in this paper is different to the Bayesian models considered for the increments $D_{i,j} = C_{i,j} - C_{i,j-1}$ in Verrall [18] and [19]

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

In the sequel it is useful to define for $j = 0, \ldots, J$

$$\mathcal{B}_j = \{C_{i,k};\ i + k \leq I,\ k \leq j\} \subset \mathcal{D}_I, \tag{2.12}$$

which is the set of observations up to development period j at time I. It is convenient to switch from the random variables $C_{i,j}$ to the random variables $Y_{i,j}$ (individual development factors) defined by

$$Y_{i,j} = \frac{C_{i,j+1}}{C_{i,j}}. \tag{2.13}$$

Without loss of generality we assume in (2.13) that $C_{i,j} > 0$. If $C_{i,j} = 0$ then the process stops, which is also the case in Mack's model, because then $E[C_{i,j+1} | C_{i,j} = 0] = 0$ and $\mathrm{Var}[C_{i,j+1} | C_{i,j} = 0] = 0$. Analogously, for $j = 0, 1, \ldots, J-1$, we denote by

$$\mathbf{Y}_j = (Y_{0,j}, Y_{1,j}, \ldots, Y_{I-j-1,j})',$$

the column vectors of the observed Y-trapezoid and by

$$\mathbf{y}_j = (y_{0,j}, y_{1,j}, \ldots, y_{I-j-1,j})'$$

a realization of $\mathbf{Y}_j$. The chain ladder assumptions (2.10) and (2.11) are then equivalent to

$$E[Y_{i,j} | C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = f_j, \tag{2.14}$$

$$\mathrm{Var}[Y_{i,j} | C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = \frac{\sigma_j^2}{C_{i,j}}. \tag{2.15}$$

In the chain ladder methodology and in the underlying stochastic chain ladder model of Mack, only the individual data of a specific claims development triangle or claims development trapezoid are considered and modelled. In this paper we also want to make use of prior information or of portfolio information from other "similar" risks, from which we can possibly learn something about the unknown claims development pattern of the considered specific claims data. To do this, we have to consider the chain ladder methodology in a Bayesian set up. This is described in the next section.

3. Bayes Chain Ladder

In the Bayesian chain ladder set up, the unknown chain ladder factors f_j, $j = 0, 1, \ldots, J-1$, are assumed to be realizations of independent positive, real valued random variables F_j. We denote by

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

$$\mathbf{F} = (F_0, F_1, \ldots, F_{J-1})'$$

the random vector of the F_j's and by

$$\mathbf{f} = (f_0, f_1, \ldots, f_{J-1})'$$

a realization of **F**. In the Bayes chain ladder model it is assumed that conditionally, given **F**, the chain ladder Model Assumptions 2.1 are fulfilled.

Model Assumptions 3.1. (Bayes chain ladder)

B1 *Conditionally, given* **F**, *the random variables $C_{i,j}$ belonging to different accident years $i \in \{0, 1, \ldots, I\}$ are independent.*

B2 *Conditionally, given* **F** *and $\{C_{i,0}, C_{i,1}, \ldots, C_{i,j}\}$, the conditional distribution of $Y_{i,j}$ only depends on F_j and $C_{i,j}$, and it holds that*

$$E[Y_{i,j} \mid \mathbf{F}, C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = F_j, \tag{3.1}$$

$$\mathrm{Var}[Y_{i,j} \mid \mathbf{F}, C_{i,0}, C_{i,1}, \ldots, C_{i,j}] = \frac{\sigma_j^2(F_j)}{C_{i,j}}. \tag{3.2}$$

B3 $\{F_0, F_1, \ldots, F_{J-1}\}$ *are independent and positive.*

Remarks:

- The conditional expected value of $Y_{i,j}$, given **F** and $\{C_{i,0}, C_{i,1}, \ldots, C_{i,j}\}$, depends only on the unknown chain ladder factor F_j and not on the chain ladder factors F_k of other development years $k \neq j$.
- In (3.2) $C_{i,j}$ plays the role of a weight function, i.e. the conditional variance of $Y_{i,j}$, given **F** and $\{C_{i,0}, C_{i,1}, \ldots, C_{i,j}\}$, is inversely proportional to $C_{i,j}$. Note that the nominator of (3.2) may depend on F_j.
- Of course, the unconditional distribution of **F** does not depend on $\mathcal{D}_I$. The distributions of the F_j's are often called structural function.
- We define

$$\widehat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}},$$

which is the estimator of the chain ladder factor f_j in the classical chain ladder model (see (2.7)). Then it follows from Model Assumptions 3.1 that

$$E\left[\widehat{F}_j \,\middle|\, \mathbf{F}, \mathcal{B}_j\right] = F_j, \tag{3.3}$$

$$\mathrm{Var}\left[\widehat{F}_j \,\middle|\, \mathbf{F}, \mathcal{B}_j\right] = \frac{\sigma_j^2(F_j)}{S_j^{[I-j-1]}}. \tag{3.4}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

- Conditionally, given $\mathbf{F}$, $\{C_{i,j} : j = 0, 1, \ldots, J\}$ possess the Markov property, i.e. the conditional distribution of $C_{i,j+1}$ given $\{C_{i,k} : k = 0, 1, \ldots, j\}$ depends only on the last observation $C_{i,j}$ and not on the observations $C_{i,k}$ for $k < j$. This is a slightly stronger assumption than assumption $M2$ of Mack (Model Assumptions 2.1), where only the conditional first and second moments and not the whole conditional distribution depends on the last observation $C_{i,j}$.
- Conditionally, given $\mathbf{F}$,

$$\{Y_{i,j} : j = 0, 1, \ldots, J-1\} \text{ are uncorrelated and} \tag{3.5}$$

$$Y_{i,j} \text{ and } Y_{k,l} \text{ are independent for } i \neq k. \tag{3.6}$$

  (3.5) is a well known result from Mack [10]. Note however, that the $Y_{i,j}$ in (3.5) are only uncorrelated but not independent (see Mack et al. [13]).
- It is sometimes convenient to consider the increments

$$D_{i,j+1} = C_{i,j+1} - C_{i,j}$$

  and to define the incremental chain ladder factors $\tilde{F}_j$ and the corresponding incremental observations $\tilde{Y}_{i,j}$ by

$$\tilde{Y}_{i,j} = \frac{D_{i,j+1}}{C_{i,j}} = Y_{i,j} - 1, \tag{3.7}$$

$$\tilde{F}_j = F_j - 1. \tag{3.8}$$

  The Bayes chain ladder conditions can also be written in terms of $\tilde{Y}_{i,j}$ and $\tilde{F}_j$ instead of $Y_{i,j}$ and F_j and are exactly the same as Model Assumptions 3.1 if we replace $Y_{i,j}$, F_j and $\mathbf{F}$ by $\tilde{Y}_{i,j}$, $\tilde{F}_j$ and $\tilde{\mathbf{F}}$, respectively.
- Our goal is to find best predictors of $C_{i,j}$ for $i + j > I$, given the observations $\mathcal{D}_I$.

In the Bayesian framework it is assumed that the distribution of the data $(Y_{i,j})$ given $\mathbf{F}$ as well as the distributions of the F_j's are specified. Then Bayes' theorem allows for the calculation of the posterior distribution of $\mathbf{F}$ given the data. The next theorem gives the result for this calculation. In practice, the conditional distribution of the data given $\mathbf{F}$ as well as the prior distributions of the F_j's are mostly unknown. What can be done in such cases is the subject of Section 4.

Theorem 3.2. *Under Model Assumptions 3.1 it holds that a posteriori, given the observations $\mathcal{D}_I$, the random variables $F_0, F_1, \ldots, F_{J-1}$ are independent with posterior distribution given by* (3.11).

Remark:

- Because of (3.7), Theorem 3.2 holds also true for $\tilde{F}_0, \tilde{F}_1, \ldots, \tilde{F}_{J-1}$.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Proof of Theorem 3.2.

To simplify notation, we use in the remainder of this section the following terminology: for $j = 0, 1, \ldots, J-1$, the distribution functions of the random variables F_j are denoted by $U(f_j)$, i.e. we use the same capital U for different distribution functions and the argument f_j in $U(f_j)$ says that it is the distribution function of F_j. Analogously, we denote the conditional distribution of the data, given $F_j = f_j$ or $\mathbf{F} = \mathbf{f}$, by $F_{f_i}(.)$ and $F_{\mathbf{f}}(.)$, respectively. For instance, $F_{f_j}(y_{i,j} | \mathcal{B}_j)$ is the conditional distribution of $Y_{i,j}$, given $F_j = f_j$ and given $\mathcal{B}_j$.

From Model Assumptions 3.1 follows that

$$dF_{\mathbf{f}}(\mathbf{y}_0, \ldots, \mathbf{y}_{J-1} | \mathcal{B}_0) = \prod_{j=0}^{J-1} \prod_{i=0}^{I-j-1} dF_{f_j}(y_{i,j} | C_{i,j}), \tag{3.9}$$

where $C_{i,j} = y_{i,j-1} C_{i,j-1}$ for $j \geq 1$. For the joint posterior distribution of $\mathbf{F}$ given the observations $\mathcal{D}_I$ we find

$$dU(f_0, \ldots, f_{J-1} | \mathcal{D}_I) \propto \prod_{j=0}^{J-1} \left\{ \prod_{i=0}^{I-j-1} dF_{f_j}\left(y_{i,j} \middle| C_{i,j}\right) dU\left(f_j\right) \right\} \tag{3.10}$$

$$\propto \prod_{j=0}^{J-1} dU\left(f_j \middle| \mathcal{D}_I\right), \tag{3.11}$$

which completes the proof of the theorem. □

Remark:

- Note that the conditional distribution of F_j, given $\mathcal{D}_I$, depends only on $\mathbf{Y}_j$ and $\mathbf{C}_j$, where $C_{i,j}$, $i = 0, \ldots, I-j$ play the role of weights because of the variance condition $\mathrm{Var}[Y_{i,j} | C_{i,0}, \ldots, C_{i,j}, F_j] = \sigma_j^2(F_j)/C_{i,j}$. Indeed, the random variables $Y_{i,j}$ are the only ones in the Y-trapezoid containing information on F_j.

Next we derive the Bayes predictor of the ultimate claim $C_{i,J}$.

Definition 3.3. *A predictor $\widehat{Z}$ of some random variable Z is said to be better or equal than a predictor $\tilde{Z}$ if*

$$E\left[\left(\widehat{Z} - Z\right)^2\right] \leq E\left[\left(\tilde{Z} - Z\right)^2\right]. \tag{3.12}$$

Definition 3.3 means that we use the expected quadratic loss as optimization criterion.

The following result is a well known result from Bayesian statistics.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Theorem 3.4. *Let Z be an unknown random variable and* $\mathbf{X}$ *a random vector of* observations. Then the best predictor of Z based on X is

$$Z^{Bayes} = E[Z \mid \mathbf{X}]. \tag{3.13}$$

Remark:

- Z^{Bayes} also minimizes the conditional quadratic loss, i.e.

$$Z^{Bayes} = \underset{\widehat{Z}}{\arg\min}\, E\left[\left(\widehat{Z} - Z\right)^2 \middle| \mathbf{X}\right]. \tag{3.14}$$

Let $\widehat{C}_{i,J}$ be a predictor of the ultimate claim $C_{i,J}$ based on the observations $\mathcal{D}_I$.

Definition 3.5. *The conditional mean square error of prediction of* $\widehat{C}_{i,J}$ *is defined by*

$$mse\left(\widehat{C}_{i,J}\right) = E\left[\left(\widehat{C}_{i,J} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right]. \tag{3.15}$$

Denote by

$$\widehat{R}_i = \widehat{C}_{i,J} - C_{i,I-i} \tag{3.16}$$

the corresponding claims reserves estimate. Then, note that

$$mse\left(\widehat{C}_{i,J}\right) = mse\left(\widehat{R}_i\right) = E\left[\left(\widehat{R}_i - R_i\right)^2 \middle| \mathcal{D}_I\right]. \tag{3.17}$$

From (3.14) follows that

$$C_{i,J}^{Bayes} = E[C_{i,J} \mid \mathcal{D}_I] \tag{3.18}$$

is the best estimator minimizing the conditional mean square error of prediction (3.15).

Theorem 3.6. *Under Model Assumptions 3.1 we have*

$$C_{i,J}^{Bayes} = C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes} \tag{3.19}$$

for $J > I - i$, where F_j^{Bayes} denotes the Bayes estimator of F_j.

Remarks:

- The corresponding claims reserves estimate is given by

$$R_i^{Bayes} = C_{i,J}^{Bayes} - C_{i,I-i}.$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

- If we consider the incremental chain ladder factors $\tilde{F}_j$, we obtain from (3.7), that

$$\tilde{F}_j^{Bayes} = F_j^{Bayes} - 1,$$

and hence

$$C_{i,J}^{Bayes} = C_{i,I-i} \prod_{j=I-i}^{J-1} \left(\tilde{F}_j^{Bayes} + 1 \right).$$

Proof of Theorem 3.6.

From the posterior independency of the F_j, given $\mathcal{D}_I$, (see Theorem 3.2) follows that $Y_{i,j}$, $j = I-i, \ldots, J-1$, are also conditionally uncorrelated. Thus we obtain

$$\begin{aligned}
C_{i,J}^{Bayes} &= E\left[C_{i,I-i} \prod_{j=I-i}^{J-1} Y_{i,j} \middle| \mathcal{D}_I \right] \\
&= C_{i,I-i} \prod_{j=I-i}^{J-1} E\left[Y_{i,j} \middle| \mathcal{D}_I \right] \\
&= C_{i,I-i} \prod_{j=I-i}^{J-1} E\left\{ E\left[Y_{i,j} \middle| \mathbf{F}, \mathcal{D}_I \right] \middle| \mathcal{D}_I \right\} \\
&= C_{i,I-i} \prod_{j=I-i}^{J-1} E\left[F_j \middle| \mathcal{D}_I \right] \\
&= C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes}.
\end{aligned}$$

□

Next we want to find a formula for the mean square error of prediction of R_i^{Bayes}, which is the same as the mean square error of prediction of $C_{i,J}^{Bayes}$. Because of the general property

$$\mathrm{Var}[X] = E[\mathrm{Var}[X|Y]] + \mathrm{Var}[E[X|Y]]$$

it holds that

$$\begin{aligned}
mse\left(C_{i,J}^{Bayes}\right) &= E\left[\left(C_{i,J}^{Bayes} - C_{i,J} \right)^2 \middle| \mathcal{D}_I \right] \\
&= E\left[\mathrm{Var}\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I \right] \middle| \mathcal{D}_I \right] \\
&\quad + E\left[\left(C_{i,J}^{Bayes} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I \right] \right)^2 \middle| \mathcal{D}_I \right].
\end{aligned} \tag{3.20}$$

In the classical approach of Mack [10] the first term corresponds to the process error and the second to the estimation error. Here, this picture is not so clear

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

any more. Since **F** is a random vector, the first term is some kind of an "average" process error (averaged over the set of possible values of **F**) and the second term is some kind of an "average" estimation error.

$$\begin{aligned}\operatorname{Var}\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] &= E\left[\operatorname{Var}\left(C_{i,J} \middle| \mathbf{F}, C_{i,J-1}\right) \middle| \mathbf{F}, \mathcal{D}_I\right] \\ &\quad + \operatorname{Var}\left(E\left[C_{i,J} \middle| \mathbf{F}, C_{i,J-1}\right] \middle| \mathbf{F}, \mathcal{D}_I\right) \\ &= C_{i,I-i}\, \sigma^2(F_{J-1}) \prod_{j=I-i}^{J-2} F_j + F_{J-1}^2 \operatorname{Var}\left[C_{i,J-1} \middle| \mathbf{F}, \mathcal{D}_I\right].\end{aligned} \tag{3.21}$$

By iterating (3.21) we obtain

$$\operatorname{Var}\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] = C_{i,I-i} \sum_{k=I-i}^{J-1} F_{I-i} \cdots F_{k-1}\, \sigma^2(F_k)\, F_{k+1}^2 \cdots F_{J-1}^2 . \tag{3.22}$$

Formula (3.22) is the same as the formula found by Mack [10], which is not surprising, since conditionally on **F**, the chain ladder model assumptions of Mack are fulfilled. The next step however differs, because the F_j's are now random. From (3.22) and since the F_j are conditionally independent, given $\mathcal{D}_I$, we obtain for the "average" process error

$$\begin{aligned}&E\left[\operatorname{Var}\left(C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right) \middle| \mathcal{D}_I\right] = \\ &\qquad C_{i,I-i} \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes}\, E\left[\sigma^2(F_k) \middle| \mathcal{D}_I\right] \prod_{n=k+1}^{J-1} E\left[F_n^2 \middle| \mathcal{D}_I\right] \right\}.\end{aligned} \tag{3.23}$$

The "average" estimation error of $C_{i,J}^{Bayes}$ is given by (using the posterior independence again)

$$\begin{aligned}E\left[\left(C_{i,J}^{Bayes} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right] &= C_{i,I-i}^2\, E\left[\left\{ \prod_{j=I-i}^{J-1} E\left[F_j \middle| \mathcal{D}_I\right] - \prod_{j=I-i}^{J-1} F_j \right\}^2 \middle| \mathcal{D}_I\right] \\ &= C_{i,I-i}^2 \operatorname{Var}\left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right).\end{aligned} \tag{3.24}$$

From (3.20), (3.23) and (3.24) follows immediately the following result:

Theorem 3.7. *The conditional mean square error of prediction of the Bayesian claims reserves of accident year i is given by*

$$\begin{aligned}mse\left(R_i^{Bayes}\right) &= E\left[\left(C_{i,J}^{Bayes} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \\ &= C_{i,I-i}\, \Gamma_{I-i} + C_{i,I-i}^2\, \Delta_{I-i}^B,\end{aligned} \tag{3.25}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

where

$$\Gamma_{I-i} = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes} E\left[\sigma^2(F_k) \middle| \mathcal{D}_I\right] \prod_{n=k+1}^{J-1} E\left[F_n^2 \middle| \mathcal{D}_I\right] \right\}, \tag{3.26}$$

$$\Delta_{I-i}^{B} = \mathrm{Var}\left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I \right). \tag{3.27}$$

4. Credibility for Chain Ladder

In the Bayesian set up, the best predictor of the ultimate claim $C_{i,J}$ is

$$C_{i,J}^{Bayes} = C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Bayes}. \tag{4.1}$$

However, to calculate F_j^{Bayes} one needs to know the distributions of the F_j as well as the conditional distributions of the $C_{i,j}$, given $\mathbf{F}$. These distributions are usually unknown in the insurance practice. The advantage of credibility theory is that one needs to know only the first and second moments. It is assumed that these first and second moments exist and are finite for all considered random variables. Given a portfolio of similar risks, these moments can be estimated from the portfolio data. For the results of credibility theory used in this paper we refer the reader to the literature, e.g. to the book by Bühlmann and Gisler [3].

By chain ladder credibility we mean that we replace the F_j^{Bayes} in (4.1) by credibility estimators F_j^{Cred}.

Definition 4.1. *The credibility based predictor of the ultimate claim $C_{i,J}$ given $\mathcal{D}_I$ is defined by*

$$C_{i,J}^{(Cred)} = C_{i,I-i} \prod_{j=I-i}^{J-1} F_j^{Cred}. \tag{4.2}$$

Remarks:

- Note that we have put the superscript *Cred* into brackets and that we call $C_i^{(Cred)}$ a *credibility based estimator* and not a credibility estimator. By definition a credibility estimator would be a linear function of the observations. However, given the multiplicative structure in the chain ladder methodology, it would not make sense to restrict to linear estimators of $C_{i,J}$.
- The corresponding reserve estimate is defined by

$$R_i^{(Cred)} = C_{i,J}^{(Cred)} - C_{i,I-i}.$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Credibility estimators based on some statistic $\mathbf{X}$ are best estimators which are a linear function of the entries of $\mathbf{X}$. For estimating F_j we base our estimator on the observations $Y_{i,j}$, $i = 0, \ldots, I - j - 1$, since these are the only observations of the Y-trapezoid containing information on F_j.

Definition 4.2.

$$F_j^{Cred} = \underset{\left\{\widehat{F}_j \,:\, \widehat{F}_j = a_0^{(j)} + \sum_{i=0}^{I-j-1} a_i^{(j)} Y_{i,j}\right\}}{\arg\min} E\left[\left(\widehat{F}_j - F_j\right)^2 \middle| \mathcal{B}_j\right]. \tag{4.3}$$

In other words, F_j^{Cred} is defined as the best estimator within $\{\widehat{F}_j : \widehat{F}_j = a_0^{(j)} + \sum_{i=0}^{I-j-1} a_i^{(j)} Y_{i,j}\}$ minimizing square error function $E[(\widehat{F}_j - F_j)^2 | \mathcal{B}_j]$.

Theorem 4.3. (Credibility estimator)

i) The credibility estimators for the chain ladder factors F_j are given by

$$F_j^{Cred} = \alpha_j \widehat{F}_j + (1 - \alpha_j) f_j, \tag{4.4}$$

where

$$\widehat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}, \tag{4.5}$$

$$\alpha_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]} + \frac{\sigma_j^2}{\tau_j^2}}, \tag{4.6}$$

and the structural parameters are given by

$$f_j = E[F_j], \tag{4.7}$$

$$\sigma_j^2 = E[\sigma_j^2(F_j)], \text{ where } \sigma_j^2(F_j) \text{ is defined in } (3.2), \tag{4.8}$$

$$\tau_j^2 = \mathrm{Var}[F_j]. \tag{4.9}$$

ii) The conditional mean square error of prediction of F_j^{Cred} is

$$mse(F_j^{Cred}) = E[(F_j^{Cred} - F_j)^2 | \mathcal{B}_j] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} = (1 - \alpha_j)\tau_j^2. \tag{4.10}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Remark:

- Note that $\widehat{F}_j$ is the estimate of the development factor f_j in the "classical" chain ladder model. That is, (4.4) is a credibility weighted average between the classical chain ladder estimator $\widehat{F}_j$ and the a priori expected value f_j (expert opinion or market experience).

Proof of the theorem:

Conditionally on $\mathcal{B}_j$, the random variables $Y_{i,j}$, $i = 0, 1, \ldots, I-j-1$, fulfil the assumptions of the Bühlmann-Straub model (see Section 4.2 in Bühlmann and Gisler [3]). Then (4.4) is the well known credibility estimator in the Bühlmann-Straub model and the formula of its mean square error is also found, for example, in Bühlmann and Gisler [3], Chapter 4. □

The credibility estimator (4.4) depends on the structural parameters f_j, σ_j^2 and τ_j^2. These structural parameters can be estimated from portfolio data by using standard estimators (see for instance Bühlmann and Gisler [3], Section 4.8).

For the conditional mean square error of prediction we obtain similar to (3.20)

$$\begin{aligned} mse\left(R_i^{(Cred)}\right) &= mse\left(C_{i,J}^{(Cred)}\right) = E\left[\left(C_{i,J}^{(Cred)} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \\ &= E\left[\mathrm{Var}\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] \middle| \mathcal{D}_I\right] + E\left[\left(C_{i,J}^{(Cred)} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right]. \end{aligned}$$

The first summand, the "average" process error, remains unchanged and is the same as in Theorem 3.7. For the "average" estimation error we obtain

$$E\left[\left(C_{i,J}^{(Cred)} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right] = C_{i,I-i}^2 \, E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right]. \tag{4.11}$$

Henceforth, we have

$$mse\left(R_i^{(Cred)}\right) = C_{i,I-i} \, \Gamma_{I-i} + C_{i,I-i}^2 \, \Delta_{I-i}^C, \tag{4.12}$$

where

$$\Gamma_{I-i} = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes} \, E\left[\sigma^2(F_k) \middle| \mathcal{D}_I\right] \prod_{n=k+1}^{J-1} E\left[F_n^2 \middle| \mathcal{D}_I\right] \right\}, \tag{4.13}$$

$$\Delta_{I-i}^C = E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right]. \tag{4.14}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Note that these two terms, in general, cannot be calculated explicitely. Therefore, to find an estimator for the conditional mean square error of prediction we make the following approximations in (4.13) and (4.14):

$$F_j^{Bayes} \simeq F_j^{Cred}, \tag{4.15}$$

$$E\left[\left(F_j - F_j^{Bayes}\right)^2 \middle| \mathcal{D}_I\right] \simeq E\left[\left(F_j - F_j^{Cred}\right)^2 \middle| \mathcal{B}_j\right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}}, \tag{4.16}$$

where $\simeq$ means that the equation is not exactly but only approximately fulfilled. Then we get

$$E\left[F_j^2 \middle| \mathcal{D}_I\right] = E\left[\left(F_j - F_j^{Bayes}\right)^2 \middle| \mathcal{D}_I\right]^2 + \left(F_j^{Bayes}\right)^2 \simeq \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} + \left(F_j^{Cred}\right)^2.$$

and using the posterior independence (see Theorem 3.2)

$$\begin{aligned} E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right] &\simeq E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Bayes} - \prod_{j=I-i}^{J-1} F_j\right)^2 \middle| \mathcal{D}_I\right] \\ = \operatorname{Var}\left(\prod_{j=I-i}^{J-1} F_j \middle| \mathcal{D}_I\right) &= \prod_{j=I-i}^{J-1} E\left[F_j^2 \middle| \mathcal{D}_I\right] - \prod_{j=I-i}^{J-1} \left(E\left[F_j \middle| \mathcal{D}_I\right]\right)^2 \\ &\simeq \prod_{j=I-i}^{J-1} \left(\alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} + \left(F_j^{Cred}\right)^2\right) - \prod_{j=I-i}^{J-1} \left(F_j^{Cred}\right)^2. \end{aligned}$$

Thus we have found the following approximation for the conditional mean square error of prediction of $R_i^{(Cred)}$:

Theorem 4.4

$$mse\left(R_i^{(Cred)}\right) \cong C_{i,I-i}\,\Gamma_{I-i}^* + C_{i,I-i}^2\,\Delta_{I-i}^*, \tag{4.17}$$

where

$$\Gamma_{I=i}^* = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Cred} \sigma_k^2 \prod_{n=k+1}^{J-1} \left(\left(F_n^{Cred}\right)^2 + \alpha_n \frac{\sigma_n^2}{S_n^{[I-n-1]}} \right) \right\}, \tag{4.18}$$

$$\Delta_{I-i}^* = \prod_{j=I-i}^{J-1} \left(\left(F_j^{Cred}\right)^2 + \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \left(F_j^{Cred}\right)^2. \tag{4.19}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Remark:

- By replacing in (4.18) and (4.19) σ_j^2 and the variance components σ_j^2 and τ_j^2 in α_j by appropriate estimates, we obtain from Theorem 4.4 an estimator for the conditional mean square error of prediction of $R_i^{(Cred)}$.

Often the conditional mean square error of prediction or the conditional prediction standard error (= square root of the conditional mean square error of prediction) of the total claims reserves is of interest too. Denote the aggregate claims reserves by

$$R^{(Cred)} = \sum_i R_i^{(Cred)}, \tag{4.20}$$

where $R_i^{(Cred)} = 0$ for those i for which $I - i \geq J$.

$$\begin{aligned} mse\left(R^{(Cred)}\right) &= E\left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \\ &= E\left[\operatorname{Var}\left[\sum_i C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] \middle| \mathcal{D}_I\right] + E\left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right]. \end{aligned} \tag{4.21}$$

Because of the conditional independence of the accident years we get for the first term

$$E\left[\operatorname{Var}\left[\sum_i C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] \middle| \mathcal{D}_I\right] = \sum_i E\left[\operatorname{Var}\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right] \middle| \mathcal{D}_I\right] = \sum_i C_{i,I-i}\, \Gamma_{I-i}, \tag{4.22}$$

where Γ_{I-i} is given by (4.13). Note that the average process error of $R^{(Cred)}$ is the sum of the individual process errors of $R_i^{(Cred)}$.

For the second summand, the average estimation error, we obtain

$$\begin{aligned} &E\left[\left(\sum_i C_{i,J}^{(Cred)} - \sum_i E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right] \\ &= \sum_{i=1}^{I} E\left[\left(C_{i,J}^{(Cred)} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)^2 \middle| \mathcal{D}_I\right] \\ &+ 2\sum_{i=1}^{I} \sum_{k=i+1}^{I} E\left[\left(C_{i,J}^{(Cred)} - E\left[C_{i,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right)\left(C_{k,J}^{(Cred)} - E\left[C_{k,J} \middle| \mathbf{F}, \mathcal{D}_I\right]\right) \middle| \mathcal{D}_I\right]. \end{aligned} \tag{4.23}$$

Hence, we obtain additional cross-covariance terms for which we need to find approximations. These cross terms derive from the fact that the same chain

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

ladder factors are used for different accident years. With the approximations (4.15) and (4.16) we get for the terms in the second summand

$$\begin{aligned}
&E\left[\left(C_{i,J}^{(Cred)} - E\left[C_{i,J}\,\middle|\,\mathbf{F},\mathcal{D}_I\right]\right)\left(C_{k,J}^{(Cred)} - E\left[C_{k,J}\,\middle|\,\mathbf{F},\mathcal{D}_I\right]\right)\middle|\,\mathcal{D}_I\right]\\
&= C_{i,I-i}\,C_{k,I-k}\,E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Cred} - \prod_{j=I-i}^{J-1} F_j\right)\left(\prod_{l=I-k}^{J-1} F_l^{Cred} - \prod_{l=I-k}^{J-1} F_l\right)\middle|\,\mathcal{D}_I\right]\\
&\simeq C_{i,I-i}\,C_{k,I-k}\,E\left[\left(\prod_{j=I-i}^{J-1} F_j^{Bayes} - \prod_{j=I-i}^{J-1} F_j\right)\left(\prod_{l=I-k}^{J-1} F_l^{Bayes} - \prod_{l=I-k}^{J-1} F_l\right)\middle|\,\mathcal{D}_I\right]\\
&= C_{i,I-i}\,C_{k,I-k}\,\mathrm{Cov}\left(\prod_{j=I-i}^{J-1} F_j,\ \prod_{l=I-k}^{J-1} F_l\,\middle|\,\mathcal{D}_I\right)\\
&= C_{i,I-i}\,C_{k,I-k}\prod_{j=I-k}^{I-i-1} F_j^{Bayes}\,\mathrm{Var}\left(\prod_{j=I-i}^{J-1} F_j\,\middle|\,\mathcal{D}_I\right)\\
&\simeq C_{i,I-i}\,C_{k,I-i}^{(Cred)}\,\Delta_{I-i}^{*},
\end{aligned} \tag{4.24}$$

where Δ_{I-i}^{*} is given by (4.13). In the second last equation we have used

$$\begin{aligned}
&\mathrm{Cov}\left(\prod_{j=I-i}^{J-1} F_j,\ \prod_{l=I-k}^{J-1} F_l\,\middle|\,\mathcal{D}_I\right)\\
&= E\left[\mathrm{Cov}\left(\prod_{j=I-i}^{J-1} F_j,\ \prod_{l=I-k}^{J-1} F_l\,\middle|\,\mathcal{D}_I, F_{I-k},\ldots,F_{I-i-1}\right)\middle|\,\mathcal{D}_I\right]\\
&+ \mathrm{Cov}\left(E\left[\prod_{j=I-i}^{J-1} F_j\,\middle|\,\mathcal{D}_I, F_{I-k},\ldots,F_{I-i-1}\right], E\left[\prod_{j=I-k}^{J-1} F_j\,\middle|\,\mathcal{D}_I, F_{I-k},\ldots,F_{I-i-1}\right]\middle|\,\mathcal{D}_I\right)\\
&= E\left[\prod_{l=I-k}^{i-1} F_l\ \mathrm{Var}\left(\prod_{j=I-i}^{J-1} F_j\,\middle|\,\mathcal{D}_I\right)\middle|\,\mathcal{D}_I\right] + 0.
\end{aligned}$$

Thus we have found the following result:

Corollary 4.5.

$$mse\left(R^{(Cred)}\right) \simeq \sum_i mse\left(R_i^{(Cred)}\right) + 2\sum_{i=1}^{I}\sum_{k=i+1}^{I} C_{i,I-i}\,C_{k,I-i}^{(Cred)}\,\Delta_{I-i}^{*}, \tag{4.25}$$

where $mse\left(R_i^{(Cred)}\right)$ *and* Δ_{I-i}^{*} *are as in Theorem 4.4.*

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

5. Link to classical Chain Ladder

In this section we consider the situation, where the information contained in the prior distribution is non-informative. This is the case if the variance τ_j^2 becomes very large, i.e. in the limiting case as $\tau_j^2 \to \infty$.

Corollary 5.1. *For $\tau_j^2 \to \infty$ (non-informative prior), the credibility based chain-ladder forecasts coincide with the classical chain ladder forecasts, that is*

$$F_j^{Cred} = \widehat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}} = \widehat{f}_j, \tag{5.1}$$

$$R_i^{(Cred)} = R_i^{CL}. \tag{5.2}$$

Remarks:

- Note, that $\widehat{f}_j$ are the estimates of the chain ladder factors in the classical chain ladder model (compare with (2.7)). Hence in this case the credibility based chain ladder forecasts are the same as the classical chain ladder forecasts and the resulting reserves are the same as the classical chain ladder reserves.

Proof:

The corollary follows immediately from Theorem 4.3, because $\alpha_j \to 1$ for $\tau_j^2 \to \infty$. $\square$

The next result shows how we can estimate the conditional mean square error of prediction in the limiting case of non-informative priors, which is in this case the conditional mean square error of prediction of the classical chain ladder forecast. Thus the following result gives another view on the estimation of the prediction error in the classical chain ladder method and suggests a different estimator to the ones found so far in the literature. The estimation of the mean square error of prediction has been the topic of several papers in the ASTIN Bulletin 36/2 (see Buchwalder et al. [2], Mack et al. [13], Gisler [6], Venter [17]). They all give different views how the prediction error can be estimated. As discussed in Gisler [6], we believe that the Bayesian approach is the appropriate way to look at the situation and to estimate the conditional mean square error of prediction. In the Bayesian approach the estimate is canonically given in the model assumptions.

From Corollary 5.1 and Theorem 4.4 we obtain immediately the following result:

Corollary 5.2. *The conditional mean square error of prediction of the chain ladder reserves R_i^{CL} can be estimated by*

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

$$\widehat{mse}\left(R_i^{CL}\right) = C_{i,I-i}\,\hat{\Gamma}_{I-i} + C_{i,I-i}^2\,\hat{\Delta}_{I-i}, \tag{5.3}$$

where

$$\hat{\Gamma}_{I-i} = \sum_{k=I-i}^{J-1}\left\{\prod_{m=I-i}^{k-1}\hat{f}_m\,\hat{\sigma}_k^2\prod_{n=k+1}^{J-1}\left(\hat{f}_n^2 + \frac{\hat{\sigma}_n^2}{S_n^{[I-n-1]}}\right)\right\}, \tag{5.4}$$

$$\hat{\Delta}_{I-i} = \prod_{j=I-i}^{J-1}\left(\hat{f}_j^2 + \frac{\hat{\sigma}_j^2}{S_j^{[I-j-1]}}\right) - \prod_{j=I-i}^{J-1}\hat{f}_j^2, \tag{5.5}$$

and where $\hat{\sigma}_j^2$ are appropriate estimators for σ_j^2.

Remark:

- These estimators should be compared to the ones suggested by Mack [10] and Buchwalder et al. [2]. The estimate $\hat{\Gamma}_{I-i}$ of the "average" process error is slightly bigger compared to the ones in Mack [10] and Buchwalder et al. [2]. The reason is that we also study the variability in the chain ladder factors F_j which results in the additional terms $\hat{\sigma}_n^2/S_n^{[I-n-1]}$. Note that for many practical applications $\hat{\sigma}_n^2/S_n^{[I-n-1]} \ll \hat{f}_j^2$.
- The estimate $\hat{\Delta}_{I-i}$ of the "average" estimation error is the same as the one in the so-called conditional resampling approach by Buchwalder et al. [2], but it is different from the one in Mack [10]. For a discussion and a comparison we refer to Chapter 3 in Wüthrich and Merz [21].

Finally, from Corollary 5.1, Corollary 5.2 and Corollary 4.5 we obtain

Corollary 5.3. *The conditional mean square error of prediction of the total reserve $R^{CL} = \sum_i R_i^{CL}$ can be estimated by*

$$\widehat{mse}\left(R^{CL}\right) = \sum_i \widehat{mse}\left(R_i^{CL}\right) + 2\sum_{i=1}^{I}\sum_{k=i+1}^{I} C_{i,I-i}\,C_{k,I-i}^{CL}\,\hat{\Delta}_{I-i}, \tag{5.6}$$

where $C_{k,I-i}^{CL}$ is defined in (2.6) *and where $\widehat{mse}(R_i^{CL})$ and $\hat{\Delta}_{I-i}$ are as in Corollary* 5.2.

6. Exact Credibility for Chain Ladder

The case where the Bayes estimator is of a credibility type is referred to as *exact credibility* in the literature. In this section we consider a class of models for the chain ladder method for which this is the case, i.e. where $F_j^{Bayes} = F_j^{Cred}$.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

For the Bayes estimator we have to specify the family of distributions. Because of the one to one relations (3.7) and (3.8) we can specify either the conditional distributions of the incremental claim figures $\tilde{Y}_{i,j}$ and the a priori distributions of the incremental chain ladder factors $\tilde{F}_j$ or the conditional distributions of the cumulative claim figures $Y_{i,j}$ and the a priori distributions of the F_j. The class of models and distributions considered in the following can be thought of either as the distributions of the cumulative or the incremental claim figures and the corresponding chain ladder factors. All results hold true for both situation. However, in the following, we will present the models and assumptions in terms of the cumulative figures.

The basic assumption of the class of models considered in this section is that, conditionally on $\mathbf{F}$ and $\mathcal{B}_j$, the random variables $Y_{0,j}, \ldots, Y_{I,j}$ are independent with a distribution belonging to the one-parameter exponential dispersion family and that the a priori distribution of F_j belongs to the family of the natural conjugate priors.

The exponential dispersion family is usually parameterized by the so called canonical parameter. Hence, instead of $\mathbf{F}$ with realizations $\mathbf{f}$ we consider in this section the vector $\boldsymbol{\Theta}$ of the canonical parameters Θ_j with realizations $\boldsymbol{\vartheta}$. Of course, as we will see, the two parameters are linked each to the other.

Definition 6.1. *A distribution G_ϑ is said to be of the exponential dispersion type, if it can be expressed as*

$$dG_\vartheta(x) = \exp\left[\frac{x\vartheta - b(\vartheta)}{\varphi / w} + c(x, \varphi / w)\right] d\nu(x), \quad x \in A \subset \mathbb{R}, \tag{6.1}$$

where

$\nu(.)$ is either the Lebesgue measure on $\mathbb{R}$ or the counting measure,

$\varphi \in \mathbb{R}^+$ is the dispersion parameter,

$w \in \mathbb{R}^+$ is a suitable weight,

$b(\vartheta)$ is a twice differentiable function with a unique inverse for the first derivative $b'(\vartheta)$.

If X has a distribution function of the exponential dispersion type (6.1), then it is well known from standard theory of generalized linear models (GLM) (see e.g. McCullagh and Nelder [14] or Dobson [4]) that

$$\mu_X = E[X] = b'(\vartheta), \tag{6.2}$$

$$\sigma_X^2 = \mathrm{Var}(X) = \frac{\varphi}{w} b''(\vartheta). \tag{6.3}$$

By taking the inverse in (6.2) we obtain

$$\vartheta = (b')^{-1}(\mu_X) = h(\mu_X), \tag{6.4}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

where $h(.)$ is called the *canonical link function*.

The variance can also be expressed as a function of the mean by

$$\mathrm{Var}(X) = \frac{\varphi}{w} b''(h(\mu_X)) = \frac{\varphi}{w} V(\mu_X), \tag{6.5}$$

where $V(.)$ is the so-called *variance function*.

Definition 6.2. *The class of distributions as defined in* (6.1) *is referred to as the one (real-valued) parameter exponential dispersion class*

$$\mathcal{F}_{\exp} = \{G_\vartheta : \vartheta \in M\}, \tag{6.6}$$

where M is the canonical parameter space (set of possible values of ϑ).

It contains the (sub-)families

$$\mathcal{F}_{\exp}^{b,c} \subset \mathcal{F}_{\exp} \tag{6.7}$$

specified by the specific form of $b(\cdot)$ and $c(\cdot,\cdot)$.

In the following we assume that, conditionally on $\mathbf{F}$ and $\mathcal{B}_j$, the random variables $Y_{0,j}, \ldots, Y_{I,j}$ are independent with a distribution belonging to a one-parameter exponential dispersion family $\mathcal{F}_{\exp}^{b,c}$. The following results hold true under this general condition. However, not all exponential dispersion families are suited for our problem. In particular, the random variables $Y_{i,j}$ need to be non-negative. Hence, only distributions having support on $\mathbb{R}^+$ are suitable for the chain ladder situation.

A subclass of the exponential dispersion class are the class of models with variance function

$$V(\mu) = \mu^p. \tag{6.8}$$

These models are defined only for p outside the interval $0 < p < 1$. For $p \leq 0$ they have positive probability mass on the whole real line. This family includes in particular the following distributions:

- $p = 0$: Normal distribution
- $p = 1$: (Overdispersed) Poisson distribution
- $1 < p < 2$: Compound Poisson distribution with Gamma distributed claim amounts, where the shape parameter of the Gamma distribution is $\gamma = (2-p)/(p-1)$.
- $p = 2$: Gamma distribution

This family of models was also considered in Ohlsson and Johansson [16] in connection with calculating risk premiums in a multiplicative tariff structure.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

In this paper there is also given a good summary on the properties of this family. In Ohlsson and Johansson [16], this family was referred to as the family of the *Tweedie models*, whereas otherwise in the literature, the Tweedie models are often restricted to the compound Poisson case with $1 < p < 2$.

For paid chain ladder, the Tweedie models with $1 < p < 2$ for modeling the incremental payments seems to us of particular interest. It assumes, that the incremental claim payments $D_{i,j} = C_{i,j+1} - C_{i,j}$ have a compound Poisson distribution, which is often a very realistic model. In a claims reserving context this family has, for example, already been studied in Wüthrich [20].

Model Assumptions 6.3. (Exponential familiy and conjugate priors)

E1 *Conditional on* $\boldsymbol{\Theta} = (\Theta_0, \ldots, \Theta_{J-1}) = \boldsymbol{\vartheta}$ *and* $\mathcal{B}_j$, *the random variables* $Y_{0,j}, \ldots, Y_{I,j}$ *are independent with distribution of the exponential dispersion type given by* (6.1) *with specific functions* $b(\cdot)$ *and* $c(\cdot,\cdot)$, *dispersion parameter* φ_j *and weights* $w_{i,j} = C_{i,j}$.

E2 $\Theta_0, \ldots, \Theta_{J-1}$ *are independent with densities (with respect to the Lebesgue measure)*

$$u_j(\vartheta) = \exp\left[\frac{f_j^{(0)}\vartheta - b(\vartheta)}{\eta_j^2} + d\left(f_j^{(0)}, \eta_j^2\right)\right], \tag{6.9}$$

where $f_j^{(0)}$ *and* η_j^2 *are hyperparameters and* $\exp\left[d\left(f_j^{(0)}, \eta_j^2\right)\right]$ *is a normalizing factor.*

Remarks:

- From Assumption E1 follows that

$$f_{\vartheta_j}\left(y_{i,j} \middle| \mathcal{B}_j\right) = \exp\left\{\frac{y_{i,j}\,\vartheta_j - b\left(\vartheta_j\right)}{\varphi_j / C_{i,j}}\right\} a\left(y_{i,j}, \varphi_j / C_{i,j}\right). \tag{6.10}$$

- From Assumption E1 also follows that

$$E[Y_{i,j} | \Theta_j] = b'(\Theta_j) = F_j, \tag{6.11}$$

$$\operatorname{Var}(Y_{i,j} | \Theta_j, C_{i,J}) = \frac{\varphi_j}{C_{i,j}}\, b''(\Theta_j) = \frac{\varphi_j}{C_{i,j}}\, V(F_j). \tag{6.12}$$

  Hence conditionally, given F_j, the chain ladder assumption $M2$ of Mack's model (Model Assumptions 2.1) are fulfilled. The difference to the Mack assumptions is, that specific assumptions on the whole conditional distribution and not only on the first and second conditional moments are made.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

For instance for Tweedie models with $p = 1$, the $Y_{i,j}$ (or the $\tilde{Y}_{i,j}$) are assumed to be conditionally overdispersed Poisson distributed, or for the Tweedie models with $1 < p < 2$, the $Y_{i,j}$ (or the $\tilde{Y}_{i,j}$) are assumed to be conditionally compound Poisson distributed.

- The distributions of $\Theta_0, \dots, \Theta_{J-1}$ with densities (6.9) belong to the family $\mathcal{U}_{\exp}^{b}$ of the natural conjugate prior distributions to $\mathcal{F}_{\exp}^{b,c}$, which is given by

$$\mathcal{U}_{\exp}^{b} = \left\{ u_\gamma(\vartheta) : \gamma = \left(x_0, \eta^2\right) \in \mathbb{R} \times \mathbb{R}^+ \right\},$$

$$u_\gamma(\vartheta) = \exp\left[\frac{x_0 \vartheta - b(\vartheta)}{\eta^2} + d\left(x_0, \eta^2\right) \right], \; \vartheta \in M.$$

- The Gaussian model and the Gamma model studied in Gisler [6] are special cases of the exponential dispersion model defined above.
- The model is formulated in terms of individual development factors $Y_{i,j}$. We could also formulate it in terms of $C_{i,j}$. Then, for given $\boldsymbol{\vartheta}$, we would obtain a time series model similar to Murphy [15], Barnett and Zehnwirth [1] or Buchwalder et al. [2].

Theorem 6.4. *Under Model Assumptions 6.3 and if the region M is such that $u_j(\vartheta)$ disappears on the boundary of M then it holds that $E[F_j] = f_j^{(0)}$. Moreover, we have*

$$F_j^{Bayes} = F_j^{Cred} = \alpha_j \widehat{F}_j + (1 - \alpha_j) f_j^{(0)}, \tag{6.13}$$

where

$$\widehat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}}, \tag{6.14}$$

$$\alpha_j = \frac{S_j^{[I-j-1]}}{S_j^{[I-j-1]} + \frac{\sigma_j^2}{\tau_j^2}}, \tag{6.15}$$

$$\sigma_j^2 = \varphi_j \, E[b''(\Theta_j)] = \varphi_j \, E[V(F_j)], \tag{6.16}$$

$$\tau_j^2 = \mathrm{Var}[b'(\Theta_j)] = \mathrm{Var}(F_j). \tag{6.17}$$

Remarks:

- Note that F_j^{Bayes} in Theorem 6.4 coincides with the credibility estimator of Theorem 4.3. Therefore $C_{i,J}^{Bayes} = C_{i,J}^{(Cred)}$.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

- For Tweedie models with $p \geq 0$ the conditions and the result of the theorem are fulfilled for $p = 0$ (Normal-distribution), for $1 < p < 2$ (compound Poisson) and for $p = 2$ (Gamma), but not for $p > 2$ (see Ohlsson and Johansson [16]).

Proof of the theorem:

The result of the theorem follows directly from well known results in the actuarial literature. It was first proved by Jewell [8] in the case without weights. A proof for the case with weights can, for instance, be found in Bühlmann and Gisler [3]. □

Theorem 6.5. *If, in addition to the conditions of Theorem 3.4, $u_j'(\vartheta)$ disappears on the boundary of M then*

i)

$$E\left[\left(F_j^{Bayes} - F_j\right)^2 \middle| \mathcal{B}_j\right] = \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} = \left(1 - \alpha_j\right)\tau_j^2. \tag{6.18}$$

ii) The conditional mean square error of prediction of the reserve R_i^{Bayes} of accident year i is given by

$$\begin{aligned} mse\left(R_i^{Bayes}\right) &= E\left[\left(C_{i,J}^{(Cred)} - C_{i,J}\right)^2 \middle| \mathcal{D}_I\right] \\ &\simeq E\left[\left(C_{i,J}^{(Cred)} - C_{i,J}\right)^2 \middle| \mathcal{B}_{I-i}\right] \\ &= C_{i,I-i}\,\Gamma_{I-i}^* + C_{i,I-i}^2\,\Delta_{I-i}^*, \end{aligned} \tag{6.19}$$

where

$$\Gamma_{I=i}^* = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} F_m^{Bayes}\, \sigma_k^2 \prod_{n=k+1}^{J-1} \left(\left(F_n^{Bayes}\right)^2 + \alpha_n \frac{\sigma_n^2}{S_n^{[I-n-1]}} \right) \right\},$$

$$\Delta_{I-i}^* = \prod_{j=I-i}^{J-1} \left(\left(F_j^{Bayes}\right)^2 + \alpha_j \frac{\sigma_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \left(F_j^{Bayes}\right)^2.$$

iii) The conditional mean square error of prediction of the total reserve $R^{Bayes} = \sum_i R_i^{Bayes}$ is

$$mse\left(R^{Bayes}\right) \simeq \sum_i mse\left(R_i^{Bayes}\right) + 2\sum_{i=1}^{I} \sum_{k=i+1}^{I} C_{i,I-i}\, C_{k,I-i}^{Bayes}\, \Delta_{I-i}^*. \tag{6.20}$$

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

Remarks:

- Note, that Γ^*_{I-i} and Δ^*_{I-i} are the same as in Theorem 4.4.
- For the Tweedie models for $p \geq 0$ the condition and the result of Theorem 6.5 are fulfilled for $p = 0$ (Normal distribution) and for $1 < p < 2$ (compound Poisson). For $p = 2$ (Gamma-distribution) it is only fulfilled for $\eta^2 < 1$. In this case the natural conjugate prior $u_j(\vartheta)$ is again a Gamma-distribution, and $\eta^2 < 1$ means, that the shape parameter γ of this Gamma distribution is smaller than one. In many practical examples this is fulfilled, see also also Section 9.2.6 in Wüthrich-Merz [21].

Proof of the theorem:

Since F_j^{Bayes} is a credibility estimator it follows from Theorem 4.3, that (6.18) holds true, if all random variables are square integrable. Thus, it remains to prove that F_j is square integrable. From (6.9) we get

$$\begin{aligned} u_j'(\vartheta) &= \frac{1}{\eta^2}\left(f_j^{(0)} - b'(\vartheta)\right) u_j(\vartheta), \\ u_j''(\vartheta) &= \frac{1}{\eta^4}\left(f_j^{(0)} - b'(\vartheta)\right)^2 u_j(\vartheta) + \frac{1}{\eta^2} b''(\vartheta)\, u_j(\vartheta). \end{aligned}$$

Since $u_j'(\vartheta)$ disappears on the boundaries of M, we have

$$\begin{aligned} 0 &= \frac{1}{\eta^4}\int_M \left(f_j^{(0)} - b'(\vartheta)\right)^2 u_j(\vartheta)\, d\vartheta + \frac{1}{\eta^2}\int_M b''(\vartheta)\, u_j(\vartheta)\, d\vartheta \\ &= \frac{1}{\eta^4}\operatorname{Var}\left(F_j\right) + \frac{1}{\eta^2}\varphi_j\, E\left[V\left(F_j\right)\right]. \end{aligned}$$

Hence, F_j is square integrable.

The proof of (6.19) is the same as in Theorem 4.4. The proof of (6.20) is the same as the derivation of Corollary 4.5. □

7. Link to Classical Chain Ladder in the case of exact credibility

In this section we consider the same exponential dispersion family Bayes models as in Section 6. But now we look at the situation of non-informative priors, i.e. we consider the limiting case for $\tau_j^2 \to \infty$.

Corollary 7.1. *Under Model Assumptions 6.3 and if $u_j(\vartheta)$ disappears on the boundaries of M, then for $\tau_j^2 \to \infty$ (non-informative prior)*

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

$$F_j^{Bayes} = \widehat{F}_j = \frac{S_{j+1}^{[I-j-1]}}{S_j^{[I-j-1]}} = \widehat{f}_j, \tag{7.1}$$

$$R_i^{Bayes} = R_i^{CL}. \tag{7.2}$$

Remarks:

- Note, that $\widehat{f}_j$ are the estimates of the chain ladder factors in the classical chain ladder model. Hence in this case the Bayes chain ladder forecasts are the same as the classical chain ladder forecasts and the resulting Bayes reserves are equal to the classical chain ladder reserves.

Proof:

The credibility weights $\alpha_j \to 1$ for $\tau_j^2 \to \infty$. The result then follows immediately from Theorem 6.4. $\square$

From Theorem 6.5 and Theorem 7.1 we also obtain immediately the following result for the donditional mean square error of prediction:

Corollary 7.2. *Under Model Assumptions 6.3 and if $u_j(\vartheta)$ and $u_j'(\vartheta)$ disappear on the boundaries of M, then it holds, for $\tau_j^2 \to \infty$ (non-informative prior), that the conditional mean square error of prediction of the chain ladder reserves R_i^{CL} can be estimated by*

$$\widehat{mse}\left(R_i^{CL}\right) = C_{i,I-i}\,\hat{\Gamma}_{I-i} + C_{i,I-i}^2\,\hat{\Delta}_{I-i}, \tag{7.3}$$

where

$$\hat{\Gamma}_{I-i} = \sum_{k=I-i}^{J-1} \left\{ \prod_{m=I-i}^{k-1} \widehat{f}_m\, \hat{\sigma}_k^2 \prod_{n=k+1}^{J-1} \left(\widehat{f}_n^{\,2} + \frac{\hat{\sigma}_n^2}{S_n^{[I-n-1]}} \right) \right\}, \tag{7.4}$$

$$\hat{\Delta}_{I-i} = \prod_{j=I-i}^{J-1} \left(\widehat{f}_j^{\,2} + \frac{\hat{\sigma}_j^2}{S_j^{[I-j-1]}} \right) - \prod_{j=I-i}^{J-1} \widehat{f}_j^{\,2}, \tag{7.5}$$

where $\hat{\sigma}_j^2$ are appropriate estimators for σ_j^2.

Remarks:

- See also the remarks before and after Corollary 5.2.
- For Tweedie models, the result holds true for $p = 0$ (Normal) and for $1 < p < 2$, but not for $p = 2$, since η^2 has to be smaller than one for $p = 2$. In particular

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

the case $1 < p < 2$ (compound Poisson with Gamma claim severities) seems to us a fairly adequate and good model for the conditional incremental chain ladder factors $\tilde{Y}_{i,j}$ in the case of paid triangles or trapezoids.

Finally, from Theorem 7.1, Theorem 7.2 and Corollary 4.5 we obtain

Corollary 7.3. *The conditional mean square error of prediction of the total reserve* $R^{CL} = \sum_i R_i^{CL}$ *can be estimated by*

$$\widehat{mse}\left(R^{CL}\right) = \sum_i \widehat{mse}\left(R_i^{CL}\right) + 2\sum_{i=1}^{I} \sum_{k=i+1}^{I} C_{i,I-i}\, C_{k,I-i}^{CL}\, \hat{\Delta}_{I-i}. \tag{7.6}$$

where $\widehat{mse}\,(R_i^{CL})$ *and* $\hat{\Delta}_{I-i}$ *are as in Corollary 7.2 with classical chain ladder factors* $\hat{f}_j$.

8. Numerical Example

For pricing and profit-analysis of different business units it is necessary to set up the claims reserves for each of these business units (BU) separately. In Appendix A trapezoids of cumulative payments of the business line contractors all risks insurance for different BU of Winterthur Insurance Company are given (for confidentiality purposes the figures are scaled with a constant). The aim is to determine for this line of business the claims reserves for each business unit. Thus we have a portfolio of similar loss development figures, which is suited to apply the theory presented in this paper.

The following table shows the estimated values of the structural parameters f_j, σ_j^2, τ_j^2, which are needed for estimating the credibility chain ladder factors F_j^{Cred}. These parameters have been estimated by the "standard estimators", which can be found e.g. in Section 4.8 of Bühlmann and Gisler [3].

estimates of structural parameters:										
j	0	1	2	3	4	5	6	7	8	9
σ_j^2	336.53	34.74	7.83	5.93	0.43	4.34	4.25	0.24	0.10	0.15
τ_j^2	-8.9E-01	2.3E-03	8.4E-04	3.2E-05	8.2E-07	-1.6E-05	7.9E-05	-2.5E-06	-1.2E-06	2.8E-06
$\kappa_j^2=\sigma_j^2/\tau_j^2$	--	15'382	9'304	185'571	521'016	--	54'005	--	--	55'089
f_j	2.111	1.129	1.033	1.013	1.004	1.001	0.993	0.998	1.000	0.999

The next table shows the development factors F_j estimated by classical chain ladder (F_j^{CL}) and by the credibility estimators (F_j^{Cred}) of Section 4 as well as the corresponding credibility weights α_j and the "weights" $S_j^{[I-j-1]}$.

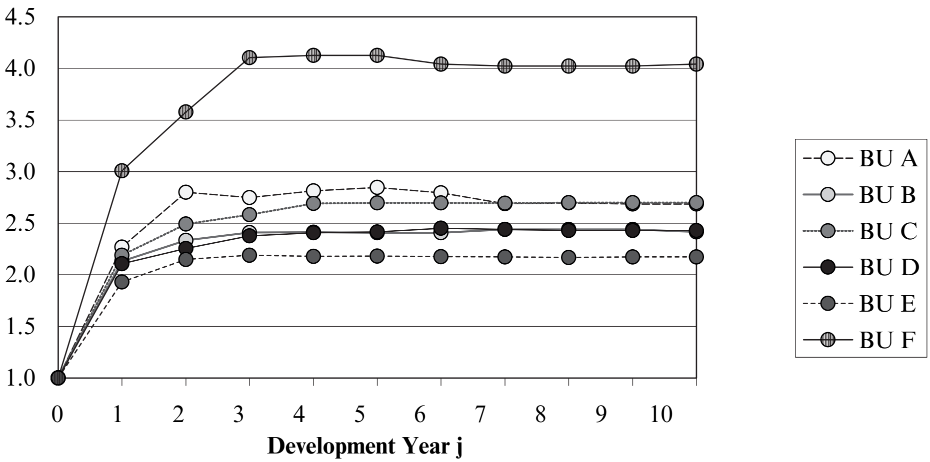
Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

	j	0	1	2	3	4	5	6	7	8	9	product
All BU	F_j	**2.111**	**1.119**	**1.031**	**1.013**	**1.004**	**1.001**	**0.993**	**0.998**	**1.000**	**0.999**	**2.452**
A	F_j^{CL}	2.270	1.233	0.982	1.024	1.012	0.981	0.962	1.003	0.996	1.000	2.686
	F_j^{cred}	2.111	1.189	0.996	1.015	1.004	1.001	0.984	0.998	1.000	0.999	2.499
	α_j	0%	57%	73%	11%	4%	0%	30%	0%	0%	23%	-
	$S_j^{[I-j-1]}$	9'203	20'482	24'863	24'028	24'075	23'769	22'678	18'500	17'067	16'026	-
B	F_j^{CL}	2.133	1.094	1.032	1.002	0.998	1.000	1.014	0.999	1.000	0.990	2.416
	F_j^{cred}	2.111	1.111	1.033	1.012	1.003	1.001	0.997	0.998	1.000	0.997	2.444
	α_j	0%	52%	65%	9%	3%	0%	20%	0%	0%	16%	-
	$S_j^{[I-j-1]}$	7'899	16'430	17'558	17'440	16'271	14'340	13'388	12'362	11'308	10'390	-
C	F_j^{CL}	2.189	1.138	1.037	1.042	1.003	1.000	0.999	1.002	1.000	1.000	2.700
	F_j^{cred}	2.111	1.134	1.036	1.016	1.004	1.001	0.994	0.998	1.000	0.999	2.509
	α_j	0%	55%	69%	10%	4%	0%	24%	0%	0%	18%	-
	$S_j^{[I-j-1]}$	8'800	18'645	20'359	20'158	19'800	18'282	17'212	14'832	13'220	11'828	-
D	F_j^{CL}	2.108	1.070	1.054	1.013	1.004	1.015	0.996	0.995	1.000	1.000	2.429
	F_j^{cred}	2.111	1.084	1.050	1.013	1.004	1.001	0.994	0.998	1.000	0.999	2.426
	α_j	0%	76%	84%	22%	9%	0%	45%	0%	0%	33%	
	$S_j^{[I-j-1]}$	23'949	48'991	50'611	52'110	49'344	46'429	43'349	35'863	33'828	26'905	
E	F_j^{CL}	1.930	1.114	1.018	0.995	1.002	0.997	0.999	0.997	1.002	1.000	2.172
	F_j^{cred}	2.111	1.119	1.021	1.010	1.004	1.001	0.995	0.998	1.000	0.999	2.429
	α_j	0%	68%	79%	16%	6%	0%	36%	0%	0%	29%	-
	$S_j^{[I-j-1]}$	17'477	32'499	35'317	34'175	32'861	31'870	30'113	27'269	24'459	22'438	-
F	F_j^{CL}	3.008	1.190	1.146	1.006	1.000	0.979	0.996	1.000	1.000	1.004	4.041
	F_j^{cred}	2.111	1.139	1.064	1.013	1.004	1.001	0.993	0.998	1.000	0.999	2.578
	α_j	0%	16%	27%	2%	1%	0%	5%	0%	0%	4%	-
	$S_j^{[I-j-1]}$	1'015	3'008	3'439	3'765	3'541	3'170	2'917	2'727	2'625	2'483	-

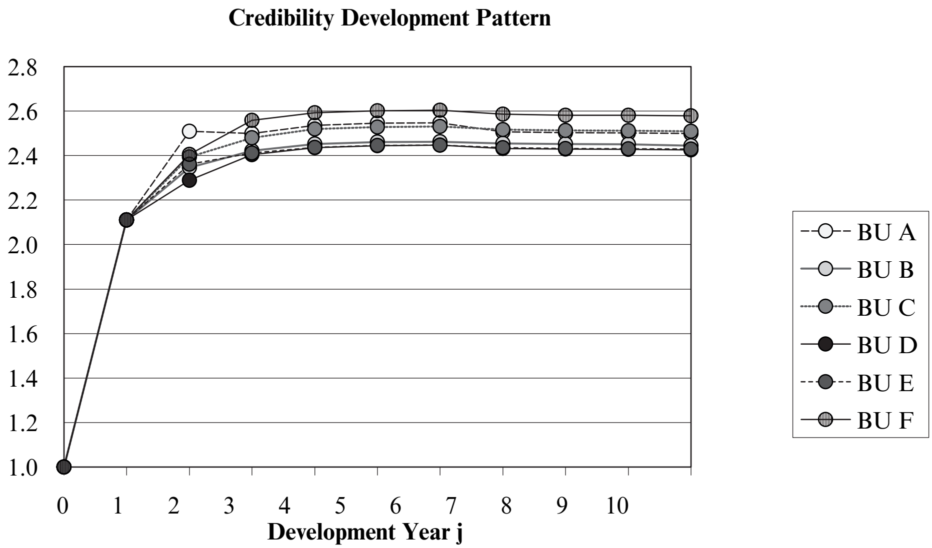
From the above table we can see that the chain ladder (CL) and the credibility (Cred) estimates can differ quite substantially (see for instance the estimate of F_0 for BU F). The estimate of the variance component τ_j^2 became negative for $j = 0, 5, 7$ and 8. Therefore F_j^{Cred} is identical to $E[F_j] = f_j$ for $j = 0, 5, 7$ and 8.

The values of the above result table can be visualized by looking at the following graphs showing the resulting loss development pattern of a "normed reference year" characterized by a payment of one in development year 0.

Chain Ladder Development Pattern

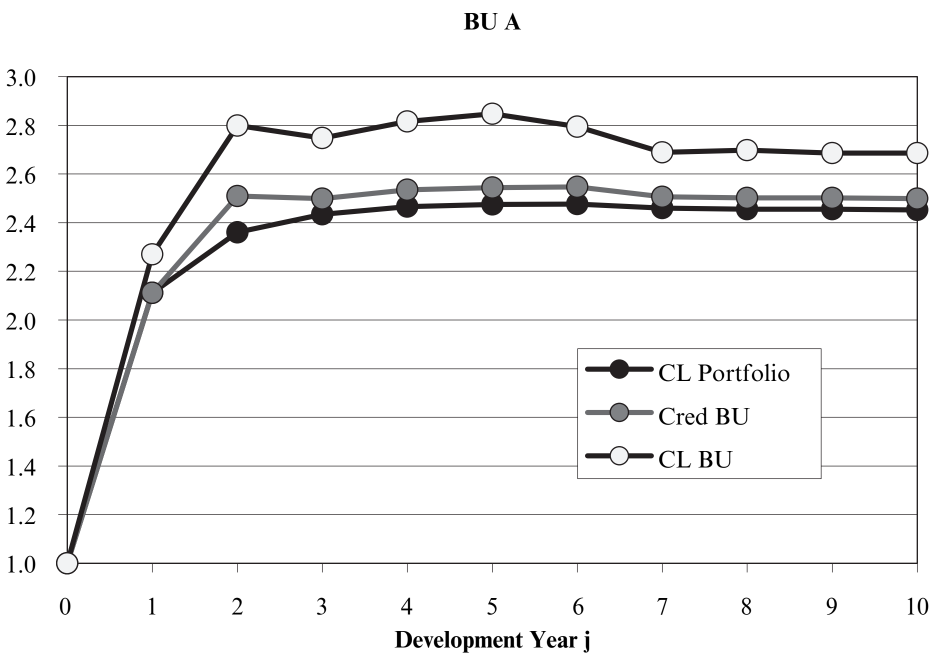


Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294



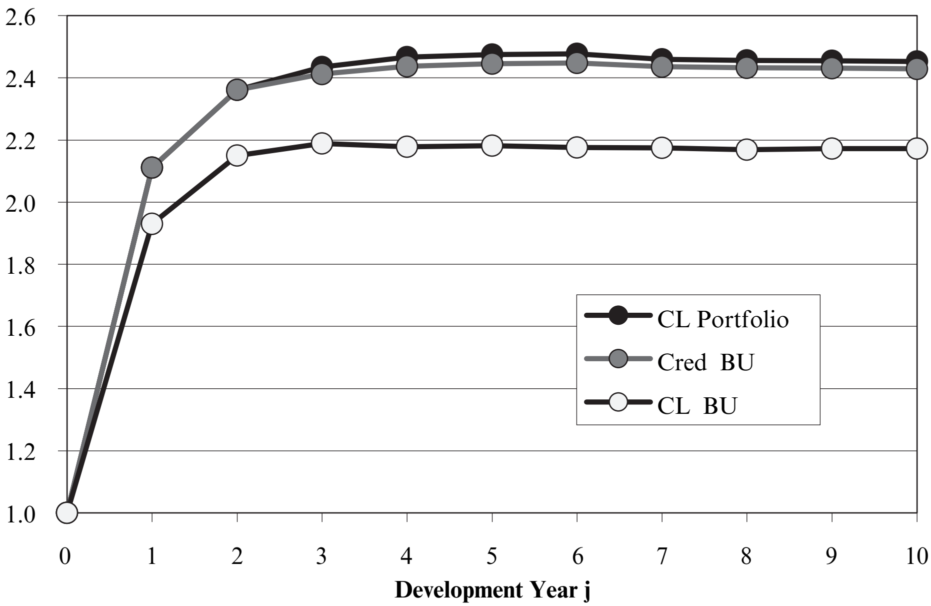
From these graphs one can see the smoothing effect on the chain ladder procedure on the estimates.

The next two graphs show for the business units A and E the same kind of development patterns. "CL Portfolio" is the one obtained by the chain ladder



Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

BU E



factors of the portfolio data (total of all BU), "CL BU" the one obtained by the chain ladder factors of the BU and "Cred BU" the one obtained by the credibility estimated chain ladder factors of the BU. One can see from these figures that Cred BU is somewhere in between "CL Portfolio" and "CL BU".

Finally the next table shows the estimated reserves and the estimates of the square root of the conditional mean square error of prediction (mse).

BU	estimated reserves		estimated $mse^{1/2}$ (with overall σ_j^2's)			estimated $mse^{1/2}$ (with indivual σ_j^2's)	
	Cred	CL	Cred	CL		CL	
				α=1	Mack	α=1	Mack
A	504	486	498	510	510	657	656
B	244	235	402	425	424	288	288
C	517	701	520	566	565	411	411
D	899	1'029	729	765	765	844	844
E	621	495	584	593	593	397	397
F	25	40	143	163	163	140	139
total	2'810	2'987	1'254.2	1'313.2	1'312.6	1'254.5	1'253.8

In total over all business units, the difference between the chain ladder reserves and the credibility reserves is 6%, and for the business units it varies between

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

4% and 38%. Hence the differences between the chain ladder and the credibility estimated reserves can be quite substantial and may have a considerable impact on the profit and loss of the individual BU.

The estimates of the mse need some further explanations. In the Bayesian or credibility model as well as in the formula of Mack [10], the formula of the estimates of the mse contain the estimates of the variance components σ_j^2 and hence the results depend on how these σ_j^2 are estimated. In the credibility based chain ladder methodology the σ_j^2 are structural parameters, which are estimated on the basis of the data of the whole portfolio by just taking the mean of the "individual" estimates obtained from the data of each business unit triangle on its own (see table below). This means that the same "overall" estimates of the σ_j^2 are used for all business units. This makes sense, since we consider a portfolio of "similar" claims triangles and since the pure random fluctuations of the "individual estimators" are very big due to the scarce data. This makes also sense for estimating the mse of the chain ladder reserves, be it by using the formula of Mack or the slightly different formula developed in this paper. The resulting figures are given in the block "estimated $mse^{1/2}$ with overall σ_j^2". The results within this block are directly comparable to each other. In the column "Cred" you find the results by applying theorem 4.4 and corollary 4.5. The results in the column "CL, $\alpha = 1$" are the mse for the CL-estimates obtained with a non-informative prior by applying Corollaries 5.2 and 5.3 or Theorem 4.4 with $\alpha = 1$ (and F_j^{cred} replaced by F_j^{CL}). Finally in the column Mack you find the results by applying Mack's formula with the overall σ_j^2. From the figures in the block "estimated $mse^{1/2}$ with overall σ_j^2" in the above table we can see, that, for each business unit, the estimated mse is smaller for the credibility reserve than for the chain ladder. However, the mse of the chain ladder reserve obtained with the new credibility formula with $\alpha = 1$ differ only little from the values obtained with the estimator suggested by Mack [10]. These findings are similar to the findings in Buchwalder et al. [2], Section 4.2 and Table 5.

For comparison purposes we have also added the estimated mse of the chain ladder reserves, if each triangle is considered on its own and if for each business unit the σ_j^2 are estimated from the data of that particular business unit (see table below). This would mean, that we do not believe, that the triangle are in some sense "similar" and that the individual estimators of the σ_j^2 are better than the "overall" estimates based on the data of the whole portfolio. However note, that the volatility in the estimates of the σ_j^2 becomes much bigger, which is then transferred to the estimates of the mse. In any case the obtained figures are not comparable with the figures in the block "estimated $mse^{1/2}$ with overall σ_j^2". But they well illustrate the dependence of the estimated mse on the estimates of the σ_j^2. Again the results obtained with the new formula ($\alpha = 1$) differ only little from the ones obtained with the Mack formula.

The chain ladder method can also be applied to the portfolio trapezoid (data from the total of all business units). By doing so, the total reserve obtained is 2'746 and the estimated $mse^{1/2}$ is 1'418. This shows once more the well known fact that the chain ladder method is not additive. Credibility could

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

estimates of "individual" σ_j^2											
BU	j=	0	1	2	3	4	5	6	7	8	9
A		418.84	87.39	6.98	1.53	1.02	7.07	18.99	0.66	0.54	0.00
B		176.15	11.25	2.65	0.38	0.71	0.00	2.66	0.00	0.00	0.87
C		58.60	6.56	9.48	28.07	0.04	0.05	0.32	0.05	0.00	0.00
D		317.92	38.22	12.79	0.61	0.72	17.28	1.43	0.56	0.00	0.00
E		134.69	14.64	6.34	4.98	0.06	0.40	2.05	0.16	0.05	0.00
F		912.98	50.36	8.73	0.03	0.00	1.25	0.03	0.00	0.00	0.06
mean		336.53	34.74	7.83	5.93	0.43	4.34	4.25	0.24	0.10	0.15

only be applied to the portfolio triangle, if the structural variance components were a priori known or fixed in a pure Bayesian way. In our paper, we have followed the empirical credibility approach and estimated the structural parameters from the portfolio data. Therefore we could not do a credibility estimate based on the trapezoid of the portfolio data.

Acknowledgment

The authors thank F. Schnelli for having carried through the calculations of the numerical example in Section 8.

References

[1] Barnett, G. and Zehnwirth, B. (2000) Best estimates for reserves. *Proc. CAS* **LXXXII**, 245-321.

[2] Buchwalder, M., Bühlmann, H., Merz, M. and Wüthrich, M.V. (2006) The mean square error of prediction in the chain ladder reserving method (Mack and Murphy revisited). *ASTIN Bulletin* **36/2**, 521-542.

[3] Bühlmann, H. and Gisler, A. (2005) *A Course in Credibility Theory and its Applications*. Universitext, Springer Verlag.

[4] Dobson, A.J. (1990) *An Introduction to Generalized Linear Models*. Chapman and Hall, London.

[5] England, P.D. and Verrall, R.J. (2002) Stochastic claims reserving in general insurance. *British Actuarial J.* **8/3**, 443-518.

[6] Gisler, A. (2006) The estimation error in the chain-ladder reserving method: a Bayesian approach. *ASTIN Bulletin* **36/2**, 554-565.

[7] Hess, K. and Schmidt, K.D. (2000) *A comparison of models for the chain-ladder method*. Dresdner Schriften zur Versicherungsmathematik, 3/2000, Technische Universität Dresden.

[8] Jewell, W.S. (1974) Credible means are exact Bayesian for exponential families. *ASTIN Bulletin* **8**, 77-90.

[9] Mack, T. (1991) A simple parametric model for rating automobile insurance or estimating IBNR claims reserves. *ASTIN Bulletin* **21/1**, 93-109.

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

[11] Mack, T. (1994) Which stochastic model is underlying the chain ladder method? *Insurance: Mathematics and Economics* **15**, 133-138.

[12] Mack, T. and Venter, G. (2000) A comparison of stochastic models that reproduce chain ladder reserve estimates. *Insurance: Mathematics and Economics* **26**, 101-107.

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

[13] 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.
[14] MCCULLAGH, P. and NELDER, J.A. (1989) *Generalized Linear Models*. Chapman and Hall, Cambridge, 2nd edition.
[15] MURPHY, D.M. (1994) Unbiased loss development factors. *Proc. CAS* **LXXXI**, 154-222.
[16] OHLSSON, E. and JOHANSSON, B. (2006) Exact credibility and Tweedie models. *ASTIN Bulletin* **36/1**, 121-133.
[17] VENTER, G. (2006) Discussion of the mean square error of prediction in the chain ladder reserving method. *ASTIN Bulletin* **36/2**, 566-572.
[18] VERRALL, R.J. (2000) An investigation into stochastic claims reserving models and chain-ladder technique. *Insurance: Mathematics and Economics* **26**, 91-99.
[19] VERRALL, R.J. (2004) A Bayesian generalized linear model for the Bornhuetter-Ferguson method of claims reserving. *North American Act. J*. **813**, 67-89.
[20] WÜTHRICH, M.V. (2003) Claims reserving using Tweedie's compound Poisson model. *ASTIN Bulletin* **33/2**, 331-346.
[21] WÜTHRICH, M.V. and MERZ, M. (2008) *Stochastic Claims Reserving Methods in Insurance*. Wiley Finance.

ALOIS GISLER
Chief Actuary Non-Life
AXA-Winterthur Insurance Company
P.O. Box 357
Email: alois.gisler@winterthur.ch

MARIO VALENTIN WÜTHRICH
Departement Mathematik
ETH Zurich
Rämistrasse 101
CH 8401 Winterthur
Email: mario.wuethrich@math.ethz.ch

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

APPENDIX

A. Loss Development Data

The following trapezoids of cumulative payements show the loss development of the line building engineering for different business units. For confidentiality reasons the data were multiplied with some constant.

BU A

	dev.year										
acc. year	0	1	2	3	4	5	6	7	8	9	10
1986	118	487	1'232	1'266	1'266	1'397	1'397	1'397	1'492	1'492	1'492
1987	124	657	863	890	914	916	941	941	941	865	865
1988	556	2'204	3'494	2'998	2'983	3'018	2'458	2'458	2'470	2'470	2'470
1989	1'646	2'351	2'492	2'507	2'612	2'612	2'608	1'755	1'755	1'755	1'755
1990	317	886	890	890	950	990	990	990	990	990	990
1991	242	919	1'218	1'224	1'229	1'249	1'249	1'249	1'249	1'249	1'249
1992	203	612	622	639	667	647	647	647	647	647	647
1993	492	1'405	1'685	1'668	1'753	1'742	1'804	1'804	1'804	1'804	1'804
1994	321	1'149	1'728	1'863	1'877	1'877	1'877	1'877	1'877	1'877	1'877
1995	609	1'109	1'283	1'294	1'253	1'255	1'255	1'255	1'255	1'255	1'255
1996	492	1'627	1'622	1'672	1'672	1'672	1'672	1'672	1'621	1'621	1'621
1997	397	793	868	889	964	964	964	964	964	964	
1998	523	1'098	1'475	1'489	1'489	1'489	1'489	1'489	1'489		
1999	1'786	2'951	3'370	3'029	3'211	3'289	3'325	3'325			
2000	241	465	536	596	652	652	652				
2001	327	622	577	583	583	583					
2002	275	520	529	529	541						
2003	89	327	378	382							
2004	295	301	396								
2005	151	406									
2006	315										

BU B

	dev.year										
acc. year	0	1	2	3	4	5	6	7	8	9	10
1986	92	442	541	541	528	528	528	528	528	528	528
1987	451	1'077	1'085	1'178	1'212	1'217	1'217	1'217	1'217	1'217	1'217
1988	404	717	834	849	849	850	850	850	850	850	850
1989	203	572	813	875	878	910	912	1'096	1'089	1'089	986
1990	352	834	1'048	1'072	1'088	1'088	1'088	1'088	1'088	1'088	1'088
1991	504	1'246	1'272	1'353	1'285	1'285	1'285	1'285	1'285	1'285	1'285
1992	509	1'008	1'061	1'061	1'061	1'071	1'071	1'071	1'071	1'071	1'071
1993	229	580	630	670	672	672	672	672	672	672	672
1994	324	815	871	859	867	777	777	777	777	777	777
1995	508	805	906	969	971	971	971	971	971	971	971
1996	354	641	833	842	842	842	842	842	842	842	842
1997	431	847	854	915	918	918	918	918	918	918	
1998	205	830	978	1'034	1'048	1'048	1'048	1'048	1'048		
1999	522	1'134	1'064	1'202	1'202	1'210	1'210	1'210			
2000	567	925	915	957	953	953	953				
2001	1'238	1'924	2'034	1'897	1'897	1'897					
2002	355	1'003	1'137	1'164	1'196						
2003	312	680	682	686							
2004	246	352	418								
2005	91	418									
2006	130										

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

BU C

	dev.year										
acc. year	0	1	2	3	4	5	6	7	8	9	10
1986	268	456	485	483	483	483	483	483	483	483	483
1987	268	520	577	579	579	579	579	579	579	579	579
1988	385	968	1'017	1'019	1'019	1'019	1'019	1'019	1'019	1'019	1'019
1989	251	742	795	931	931	931	931	931	931	931	931
1990	456	905	1'162	1'164	1'164	1'164	1'164	1'164	1'191	1'191	1'191
1991	477	1'286	1'376	1'376	1'373	1'373	1'373	1'373	1'373	1'373	1'373
1992	405	999	1'172	1'196	1'196	1'210	1'210	1'210	1'210	1'210	1'210
1993	443	932	952	965	984	992	1'012	1'012	1'012	1'012	1'012
1994	477	1'046	1'336	1'362	1'375	1'375	1'375	1'375	1'375	1'375	1'375
1995	581	1'146	1'316	1'362	1'391	1'391	1'391	1'391	1'391	1'391	1'391
1996	401	997	1'229	1'248	1'281	1'284	1'264	1'264	1'264	1'264	1'264
1997	474	778	939	1'321	1'366	1'392	1'392	1'392	1'392	1'392	
1998	649	1'420	1'707	1'709	1'709	1'709	1'709	1'638	1'638		
1999	911	1'935	2'304	2'307	2'309	2'309	2'309	2'362			
2000	508	1'054	1'101	1'071	1'071	1'071	1'071				
2001	389	790	868	909	1'569	1'569					
2002	373	998	1'091	1'155	1'201						
2003	276	853	932	948							
2004	465	820	859								
2005	343	622									
2006	254										

BU D

	dev.year										
acc. year	0	1	2	3	4	5	6	7	8	9	10
1986	330	1'022	1'066	1'086	1'094	1'094	1'094	1'094	1'094	1'094	1'094
1987	327	873	1'057	1'076	1'082	1'082	1'082	1'082	1'082	1'082	1'082
1988	304	1'137	1'234	1'460	1'475	1'588	1'586	1'586	1'586	1'586	1'586
1989	426	1'289	1'418	1'574	1'578	1'634	2'250	2'044	2'044	2'044	2'044
1990	750	2'158	2'910	3'071	3'213	3'199	3'052	3'052	3'052	3'052	3'052
1991	761	2'164	2'446	2'570	2'578	2'558	2'558	2'558	2'558	2'558	2'558
1992	1'119	2'666	2'946	3'008	3'021	3'022	3'019	3'019	3'019	3'019	3'019
1993	917	2'458	2'892	3'502	3'629	3'664	3'887	3'867	3'697	3'697	3'697
1994	905	2'014	2'459	2'466	2'554	2'554	2'554	2'540	2'540	2'540	2'540
1995	1'761	2'990	3'235	3'795	3'816	3'841	3'842	3'860	3'860	3'860	3'860
1996	824	2'063	2'378	2'368	2'384	2'368	2'373	2'373	2'373	2'373	2'373
1997	4'364	6'630	6'850	6'885	6'923	6'923	6'923	6'923	6'923	6'923	
1998	493	1'587	1'780	1'794	1'838	1'838	1'838	1'865	1'865		
1999	4'092	7'710	6'596	7'201	7'292	7'292	7'292	7'292			
2000	1'733	3'647	3'699	3'780	3'773	3'773	3'773				
2001	1'261	2'658	3'063	3'036	3'093	3'095					
2002	1'517	3'054	3'335	3'438	3'438						
2003	778	1'212	1'247	1'215							
2004	727	1'661	1'816								
2005	561	1'486									
2006	459										

BU E

	dev.year										
acc. year	0	1	2	3	4	5	6	7	8	9	10
1986	486	964	1'057	1'106	1'130	1'130	1'138	1'131	1'131	1'131	1'131
1987	867	1'669	1'643	1'717	1'720	1'724	1'724	1'724	1'724	1'724	1'724
1988	1'285	1'925	2'204	2'488	2'507	2'509	2'510	2'510	2'436	2'436	2'436
1989	395	994	1'309	1'442	1'467	1'467	1'477	1'477	1'477	1'477	1'477
1990	802	1'468	1'776	1'823	1'827	1'832	1'833	1'833	1'833	1'833	1'833
1991	966	1'967	2'628	2'743	2'294	2'338	2'358	2'358	2'358	2'358	2'358
1992	759	1'766	1'922	1'863	1'886	1'886	1'886	1'886	1'886	1'886	1'886
1993	1'136	2'139	2'219	1'921	1'931	1'944	1'947	1'867	1'867	1'867	1'867
1994	1'467	2'243	2'553	2'598	2'598	2'598	2'598	2'598	2'598	2'598	2'598
1995	1'309	2'521	2'660	2'640	2'639	2'641	2'659	2'659	2'659	2'659	2'659
1996	877	2'170	2'341	2'420	2'516	2'516	2'431	2'431	2'431	2'468	2'468
1997	1'004	1'963	2'260	2'226	2'226	2'215	2'215	2'059	2'059	2'059	
1998	1'351	2'579	2'736	2'759	2'760	2'766	2'688	2'737	2'737		
1999	906	2'341	2'667	2'655	2'655	2'650	2'650	2'824			
2000	563	1'450	1'575	1'603	1'654	1'654	1'675				
2001	417	1'006	1'034	1'049	1'049	1'050					
2002	322	836	1'046	1'123	1'143						
2003	1'047	1'656	1'689	1'779							
2004	497	843	877								
2005	1'021	1'237									
2006	302										

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294

BU F

acc. year	dev.year 0	1	2	3	4	5	6	7	8
1986	18	64	64	64	64	64	64	64	64
1987	20	73	103	153	155	155	155	155	155
1988	20	70	318	328	328	328	328	328	328
1989	88	133	133	133	133	133	133	133	133
1990	3	180	214	214	215	215	215	215	215
1991	11	79	80	82	81	81	81	81	81
1992	17	66	105	172	172	172	188	188	188
1993	73	216	218	218	218	218	218	218	218
1994	48	213	253	386	400	400	317	304	304
1995	98	153	153	158	158	158	158	158	158
1996	38	529	557	632	639	639	639	639	639
1997	42	140	141	141	141	141	141	141	141
1998	64	95	95	102	102	102	102	102	102
1999	57	144	169	178	178	178	178	178	
2000	85	178	188	186	186	186	186		
2001	212	341	357	371	371	371			
2002	56	152	187	246	246				
2003	25	44	103	178					
2004	19	137	140						
2005	25	45							
2006	7								

Downloaded from https:/www.cambridge.org/core. University of Basel Library, on 30 May 2017 at 18:51:21, subject to the Cambridge Core terms of use, available at https:/www.cambridge.org/core/terms. https://doi.org/10.1017/S0515036100015294