**Zeitschrift:** Mitteilungen / Schweizerische Aktuarvereinigung = Bulletin / Association Suisse des Actuaires = Bulletin / Swiss Association of Actuaries

**Herausgeber:** Schweizerische Aktuarvereinigung

**Band:** - (1998)

**Heft:** 2

**Artikel:** Homogeneous credibility estimators

**Autor:** Sundt, Bjørn

**DOI:** https://doi.org/10.5169/seals-551274

# B. Wissenschaftliche Mitteilungen

BJØRN SUNDT, Bergen

## Homogeneous credibility estimators

### 1 Introduction

In a book review (Sundt (1997)) the present author criticises the presentation of homogeneous credibility estimators in Dannenburg, Kaas, & Goovaerts (1996). He says in particular (p. 91) that in his opinion

> homogeneous credibility estimators make sense only in situations where the parameters of the unconditional means are allowed to vary in such a way that there exists a linear combination of the observations whose mean is equal to the mean of the estimand for all values of the parameters.

The purpose of the present paper is to elaborate more on this statement and consider the development of the theory of homogeneous credibility estimators in a historic perspective.

Homogeneous credibility estimators were first introduced by Bühlmann & Straub (1970) in a model with time-homogeneous means. The concept was generalised to a regression model by Hachemeister (1975), and in that setting it was further elaborated upon by Taylor (1977).

In Sections 2 and 3 respectively we discuss the Bühlmann-Straub and Hachemeister models.

In Section 4 we prove a general result on the relation between homogeneous and inhomogeneous credibility estimators, from which the estimators of the Bühlmann-Straub and Hachemeister models appear as special cases. This result appeared in the graduate thesis of the present author (Sundt (1978)). We apply the result to a regression model presented by Sundt (1987).

In Section 5 we finally show how both the inhomogeneous and the homogeneous credibility estimators appear as special cases of an inhomogeneous credibility estimator within a hierarchical model, generalising a result presented by Jewell (1975).

In this paper we shall always tacitly assume that all matrices that we invert, are invertible.

## 2 The Bühlmann-Straub model

2A. The following model was presented by Bühlmann & Straub (1970).
We consider $k$ independent reinsurance treaties that have been in force for $n$ years. For treaty $i$ and year $j$, let $S_{ij}$ denote the aggregate claims and $P_{ij}$ some measure of risk volume, e. g. ceded premium. We introduce the corresponding loss ratio $X_{ij} = S_{ij}/P_{ij}$. We assume that for fixed $i$, the $X_{ij}$'s are conditionally independent given an unknown random risk parameter $\Theta_i$ that characterises the treaty. It is assumed that $\Theta_1, \Theta_2, \ldots, \Theta_k$ are independent and identically distributed. Furthermore, we assume that

$$\mathrm{E}[X_{ij} \mid \Theta_i] = m(\Theta_i)\,; \qquad \mathrm{Var}[X_{ij} \mid \Theta_i] = \frac{v(\Theta_i)}{P_{ij}} \tag{2.1}$$

and introduce the structure parameters

$$\mu = \mathrm{E}m(\Theta_i)\,; \qquad \varphi = \mathrm{E}v(\Theta_i)\,; \qquad \lambda = \mathrm{Var}\,m(\Theta_i)\,. \tag{2.2}$$

For rating purposes we want to estimate $m(\Theta_l)$. As optimality criterion for estimators we use minimisation of expected quadratic loss, that is, when choosing an estimator $\dot{m}_l$, we want to make $\mathrm{E}(\dot{m}_l - m(\Theta_l))^2$ as small as possible. By the *inhomogeneous credibility estimator* $\widetilde{m}_l$ of $m(\Theta_l)$ based on the $X_{ij}$'s we mean the best estimator of the form $g_0 + \sum_{i,j} g_{ij} X_{ij}$, where $g_0$ and the $g_{ij}$'s are constants. Bühlmann (1971) showed that

$$\widetilde{m}_l = \zeta_l X_{l\cdot} + (1 - \zeta_l)\mu \tag{2.3}$$

with

$$\zeta_l = \frac{P_{l\cdot}}{P_{l\cdot} + \kappa} \tag{2.4}$$

$$P_{l\cdot} = \sum_{j=1}^{n} P_{lj}\,; \qquad X_{l\cdot} = \frac{1}{P_{l\cdot}} \sum_{j=1}^{n} P_{lj} X_{lj}\,; \qquad \kappa = \frac{\varphi}{\lambda}\,. \tag{2.5}$$

We see that $\mathrm{E}\widetilde{m}_l = \mathrm{E}m(\Theta_l)$, that is, the inhomogeneous credibility estimators imply an expected equilibrium between premiums and claims in the portfolio.
The structure parameters $\mu$, $\varphi$, and $\lambda$ would normally be unknown and have to be estimated from portfolio data. The parameter $\mu$ can be avoided by restricting the class of estimators of $m(\Theta_l)$ to estimators of the form $\dot{m}_l = \sum_{i,j} g_{ij} X_{ij}$, that is,

we let $g_0 = 0$. To obtain an expected equilibrium between premiums and claims in the portfolio, we also require that $\mathrm{E}\dot{m}_l = \mathrm{E}m(\Theta_l)$, that is,

$$\sum_{i,j} g_{ij} = 1\,. \tag{2.6}$$

The best estimator that satisfies these criteria, $\ddot{m}_l$, is called the *homogeneous credibility estimator* of $m(\Theta_l)$ based on the $X_{ij}$'s. It is shown in Bühlmann & Straub (1970) and Bühlmann (1971) that

$$\ddot{m}_l = \zeta_l X_{l\cdot} + (1-\zeta_l)\,\widehat{\mu} \tag{2.7}$$

with

$$\widehat{\mu} = \frac{\sum\limits_{i=1}^{k} \zeta_i X_{i\cdot}}{\sum\limits_{i=1}^{k} \zeta_i}\,. \tag{2.8}$$

We see that both $\widehat{\mu}$ and $\ddot{m}_l$ depend on $\varphi$ and $\lambda$. Bühlmann & Straub (1970) and Bühlmann (1971) present estimators of these parameters. The estimators can be inserted into the expressions for $\widehat{\mu}$ and $\ddot{m}_l$.

2B. Although the homogeneous credibility estimator contains a built-in estimator of $\mu$, the present author believes that one should not discard the inhomogeneous credibility estimator. Sundt (1997) argues that the homogeneous estimator is interesting only to the extent that it motivates an estimator for $\mu$. He says (p. 91),

> Let us make a parallel with life assurance based on Makeham's mortality law. This law contains three parameters. When setting premiums for a portfolio, you do not base your premiums on estimates of the Makeham parameters from your present portfolio; you use estimates found earlier from other populations. The reviewer finds that that is also a natural approach in credibility theory. One should see the credibility estimators (possibly containing unknown parameters) and parameter estimation as two separate issues.

2C. For the homogeneous credibility estimator $\ddot{m}_l$ we made the equilibrium constraint (2.6). The following theorem shows what happens if we drop that constraint.

**Theorem 2. 1.** *Let $\breve{m}_l$ be the best estimator of $m(\Theta_l)$ of the form $\sum_{ij} g_{ij} X_{ij}$. Then*

$$\breve{m}_l = \zeta_l X_{l\cdot} + (1-\zeta_l)\widehat{\mu}\,\psi \tag{2.9}$$

*with*

$$\psi = \frac{\mu^2\zeta.}{\mu^2\zeta. + \lambda} \tag{2.10}$$

$$\zeta. = \sum_{i=1}^{k} \zeta_i.$$

**Proof.** Let

$$Q = \mathrm{E}\Big(\sum_{i,j} g_{ij}X_{ij} - m(\Theta_l)\Big)^2.$$

For $r = 1,\ldots,k$; $s = 1,\ldots,n$ we have

$$\frac{dQ}{dg_{rs}} = 2\mathrm{E}\Big(\sum_{i,j} g_{ij}X_{ij} - m(\Theta_l)\Big)X_{rs}.$$

By setting these derivatives equal to zero we obtain

$$\sum_{i,j} g_{ij}\mathrm{E}X_{ij}X_{rs} = \mathrm{E}m(\Theta_l)X_{rs},$$

that is,

$$\sum_{i,j} g_{ij}\Big(\mathrm{Cov}(X_{ij},X_{rs}) + \mathrm{E}X_{ij}\mathrm{E}X_{rs}\Big) = \mathrm{Cov}(m(\Theta_l),X_{rs}) + \mathrm{E}m(\Theta_l)\mathrm{E}X_{rs}.$$

Introduction of

$$g_r. = \sum_j g_{rj}; \qquad g_{..} = \sum_{i,j} g_{ij}$$

and the Kronecker delta

$$\delta_{lr} = \begin{cases} 1 & (r = l) \\ 0 & (r \neq l) \end{cases}$$

gives

$$\frac{\varphi}{P_{rs}}g_{rs} + \lambda g_r. + \mu^2 g_{..} = \delta_{lr}\lambda + \mu^2,$$

from which we obtain

$$g_{rs} = \frac{P_{rs}}{\varphi}\left(\delta_{lr}\lambda + \mu^2(1-g_{\cdot\cdot}) - \lambda g_{r\cdot}\right). \tag{2.11}$$

Summation over $s$ gives

$$g_{r\cdot} = \frac{P_{r\cdot}}{\varphi}\left(\delta_{lr}\lambda + \mu^2(1-g_{\cdot\cdot}) - \lambda g_{r\cdot}\right),$$

which we rearrange to

$$g_{r\cdot} = \zeta_r\left(\delta_{lr} + \frac{\mu^2}{\lambda}(1-g_{\cdot\cdot})\right). \tag{2.12}$$

By summing over $r$ we obtain

$$g_{\cdot\cdot} = \zeta_l + \frac{\mu^2}{\lambda}\zeta_{\cdot}(1-g_{\cdot\cdot}),$$

and some rearranging gives

$$1 - g_{\cdot\cdot} = \frac{\lambda}{\mu^2\zeta_{\cdot}}(1-\zeta_l)\psi. \tag{2.13}$$

By insertion in (2.11) of (2.12),

$$1 - \zeta_r = \frac{\kappa}{P_{r\cdot}}\zeta_r,$$

and (2.13) successively and some rearranging we obtain

$$\begin{aligned} g_{rs} &= \frac{P_{rs}}{\varphi}(1-\zeta_r)\left(\lambda\delta_{lr} + \mu^2(1-g_{\cdot\cdot})\right) \\ &= \frac{P_{rs}}{P_{r\cdot}}\zeta_r\left(\delta_{lr} + \frac{\mu^2}{\lambda}(1-g_{\cdot\cdot})\right) \\ &= \left(\zeta_l\delta_{lr} + (1-\zeta_l)\psi\frac{\zeta_r}{\zeta_{\cdot}}\right)\frac{P_{rs}}{P_{r\cdot}}, \end{aligned}$$

from which (2.9) follows. Q.E.D.

As $\psi < 1$, we see that $\mathrm{E}\breve{m}_l < \mathrm{E}m(\Theta_l)$. In the case when the estimators are applied to set premiums, this implies that the expected premium income would be less than the expected claims. This is obviously unfortunate for a reinsurance company.
It is interesting to notice that the "shrinkage" constant $\psi$ works only on the estimator of the portfolio mean, not on the estimator of the treaty mean.
Under the mild regularity condition that $\zeta. \uparrow \infty$ when $k \uparrow \infty$, we see that $\psi \uparrow 1$ when $k \uparrow \infty$, and thus $\mathrm{E}\breve{m}_l \uparrow \mathrm{E}m(\Theta_l)$, that is, $\breve{m}_l$ is asymptotically unbiased when the number of policies increases. This condition is satisfied in particular when $P_i. \geq c > 0$ for all $i$ and some constant $c$.
We see that when $\mu = 0$, then $\breve{m}_l = \widetilde{m}_l$. This result is obvious as when $\mu = 0$, the inhomogeneous credibility estimator must be homogeneous.
When $\mu > 0$, we can rewrite (2.10) as

$$\psi = \frac{\zeta.}{\zeta. + \varrho} \tag{2.14}$$

with

$$\varrho = \frac{\lambda}{\mu^2}\,.$$

This expression has a strong resemblance with the expression (2.4) for the credibility weight $\zeta_l$; in (2.14) $\zeta.$ has taken the place of $P_l.$ as measure of risk volume.

2D. In addition to the bias, compared to the homogeneous credibility estimator, $\breve{m}_l$ has the disadvantage that it depends on the parameter $\mu$ whereas the main reason for studying homogeneous estimators seems to be to obtain a built-in estimator for $\mu$. Bühlmann & Straub (1970) also considered the following more general version of the Bühlmann-Straub model where we have a similar problem.
We leave the assumption that the $\Theta_i$'s are independent and identically distributed and generalise (2.1) and (2.2) to

$$\mathrm{E}[X_{ij} \mid \Theta_i] = m_i(\Theta_i)\,; \qquad \mathrm{Var}[X_{ij} \mid \Theta_i] = \frac{v_i(\Theta_i)}{P_{ij}}$$

$$\mu_i = \mathrm{E}m_i(\Theta_i)\,; \qquad \varphi_i = \mathrm{E}v_i(\Theta_i)\,; \qquad \lambda_i = \mathrm{Var}\, m_i(\Theta_i)\,.$$

If we define the homogeneous credibility estimator of $m_l(\Theta_l)$ as the best homogeneous linear estimator $\dot{m}_l = \sum_{i,j} g_{ij} X_{ij}$ that satisfies the unbiasedness constraint $\mathrm{E}\dot{m}_l = \mathrm{E}m(\Theta_l)$, then this constraint becomes

$$\sum_{i=1}^{k} g_i.\, \mu_i = \mu_l\,. \tag{2.15}$$

As we now have not assumed any connection between the $\mu_i$'s, they do not disappear like in (2.6). The coefficients of $\ddot{m}_l$ therefore cannot be independent of the $\mu_i$'s in the present general set-up.
It now seems natural to conclude that if a homogeneous credibility estimator should be applicable, then it has to be possible to express the mean of the estimand as a linear function of the means of the observations. We shall make this idea more precise in Section 4.

## 3 Hachemeister's regression model

In the Bühlmann-Straub model the means of the $X_{ij}$'s were assumed to be independent of time. Hachemeister (1975) allowed for dependence of time by introducing a regression assumption. We shall introduce the model somewhat more general than Hachemeister, cf. e. g. Taylor (1977).
We consider $k$ independent insurance portfolios. For portfolio $i$ we have observed the random $n_i \times 1$ vector $\mathbf{X}_i$. This portfolio depends on the unknown random risk parameter $\Theta_i$, and we assume that the $\Theta_i$'s are independent and identically distributed. Furthermore we assume that

$$\mathrm{E}[\mathbf{X}_i \mid \Theta_i] = \mathbf{Y}_i \mathbf{b}(\Theta_i)$$

and introduce the structure parameters

$$\boldsymbol{\beta} = \mathrm{E}\mathbf{b}(\Theta_i)\,; \qquad \boldsymbol{\Lambda} = \mathrm{Cov}\,\mathbf{b}(\Theta_i)\,; \qquad \boldsymbol{\Phi}_i = \mathrm{E}\,\mathrm{Cov}[\mathbf{X}_i \mid \Theta_i)]\,.$$

The quantity $\mathbf{Y}_i$ is a non-random $n_i \times q$ matrix of full rank $q \leq n_i$.
The Bühlmann-Straub model appears as a special case by letting $n_i = n$, $q = 1$, $\mathbf{Y}_i = (1, \ldots, 1)'$, and $\Phi_i = \varphi\,\mathrm{diag}(P_{i1}^{-1}, \ldots, P_{in}^{-1})$.
The inhomogeneous credibility estimator $\widetilde{m}_l$ of $m_l(\Theta_l) = \mathbf{a}_l'\mathbf{b}(\Theta_l)$, where $\mathbf{a}_l$ is a non-random $q \times 1$ vector, is given by

$$\widetilde{m}_l = \mathbf{a}_l' \left[ \mathbf{Z}_l \widehat{\mathbf{b}}_l + (\mathbf{I} - \mathbf{Z}_l)\boldsymbol{\beta} \right] \tag{3.1}$$

with

$$\widehat{\mathbf{b}}_l = \left(\mathbf{Y}_l'\boldsymbol{\Phi}_l^{-1}\mathbf{Y}_l\right)^{-1} \mathbf{Y}_l'\boldsymbol{\Phi}_l^{-1}\mathbf{X}_l$$
$$\mathbf{Z}_l = \boldsymbol{\Lambda}\mathbf{Y}_l'\boldsymbol{\Phi}_l^{-1}\mathbf{Y}_l \left(\mathbf{I} + \boldsymbol{\Lambda}\mathbf{Y}_l'\boldsymbol{\Phi}_l^{-1}\mathbf{Y}_l\right)^{-1}.$$

Hachemeister also generalised the homogeneous credibility estimator of the Bühlmann-Straub model to the regression model by defining the homogeneous credibility estimator $\breve{m}_l$ to be the best homogeneous linear estimator $\dot{m}_l = \sum_{i=1}^{k} \mathbf{g}_i'\mathbf{X}_i$ of $m_l(\Theta_l)$ that satisfies the unbiasedness constraint $\mathrm{E}\dot{m}_l = \mathrm{E}m_l(\Theta_l)$, that is,

$$\left(\sum_{i=1}^{k} \mathbf{g}_i'\mathbf{Y}_i - \mathbf{a}_l'\right)\boldsymbol{\beta} = 0\,. \tag{3.2}$$

Unfortunately the coefficients of the optimal estimator turned out to depend on $\boldsymbol{\beta}$. This is not surprising, as the unbiasedness constraint (3.2), unlike the special case (2.6), depends on $\boldsymbol{\beta}$.
The unbiasedness constraint (3.2) depends on the parameter vector $\boldsymbol{\beta}$. However, the main reason for using a homogeneous credibility estimator seemed to be to obtain a built-in estimator of $\boldsymbol{\beta}$ so that one does not need to estimate it separately. The constraint (3.2) requires that the mean of the estimator should be equal to the mean of the estimand for one particular value of $\boldsymbol{\beta}$, the real value, which we then would have to estimate separately. However, if $\boldsymbol{\beta}$ is unknown, then what we need, is the constraint that the mean of the estimator should be equal to the mean of the estimand for any possible value of $\boldsymbol{\beta}$, that is, that (3.2) should hold for any value of $\boldsymbol{\beta}$. This gives the stronger constraint

$$\sum_{i=1}^{k} \mathbf{g}_i'\mathbf{Y}_i = \mathbf{a}_l'\,, \tag{3.3}$$

which was introduced in a special case with $q = 2$ by Taylor (1975) and generalised by Taylor (1977); it was also discussed by De Vylder (1976). By defining the homogeneous credibility estimator $\ddot{m}_l$ of $m_l(\Theta_l)$ to be the best homogeneous linear estimator $\sum_{i=1}^{k} \mathbf{g}_i'\mathbf{X}_i$ that satisfies this constraint, we obtain

$$\ddot{m}_l = \mathbf{a}_l'\left[\mathbf{Z}_l\widehat{\mathbf{b}}_l + (\mathbf{I} - \mathbf{Z}_l)\widehat{\boldsymbol{\beta}}\right] \tag{3.4}$$

with

$$\widehat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{k} \mathbf{Z}_i\right)^{-1} \sum_{i=1}^{k} \mathbf{Z}_i\widehat{\mathbf{B}}_i\,. \tag{3.5}$$

Formulae (3.3)–(3.5) represent a generalisation of (2.6)–(2.8).

## 4 The general case

4A. Let us now consider the situation in more generality. We assume that we have observed the random $n \times 1$ vector $\mathbf{X}$ and wish to estimate the unknown random variable $M$; in the model of Section 3 we have $\mathbf{X} = (\mathbf{X}_1', \ldots, \mathbf{X}_k')'$ and $M = m_l(\Theta_l)$.
We shall need the following theorem; for a proof cf. e. g. Theorem 6.1 in Sundt (1993).

**Theorem 4.1.** *An inhomogeneous linear estimator $\dot{M}$ of $M$ based on $\mathbf{X}$ is an inhomogeneous credibility estimator if and only if it satisfies the normal equations*

$$\mathrm{Cov}(\dot{M} - M, \mathbf{X}') = \mathbf{0}\,; \qquad \mathrm{E}(\dot{M} - M) = 0\,.$$

*If both $\dot{M}$ and $\overset{\circ}{M}$ are inhomogeneous linear estimators satisfying these equations, then $\dot{M} = \overset{\circ}{M}$ with probability one.*

From Theorem 4.1 we obtain that the inhomogeneous credibility estimator $\widetilde{M}$ of $M$ based on $\mathbf{X}$ is given by

$$\widetilde{M} = \mathrm{Cov}(M, \mathbf{X}')(\mathrm{Cov}\,\mathbf{X})^{-1}(\mathbf{X} - \mathrm{E}\mathbf{X}) + \mathrm{E}M\,. \tag{4.1}$$

De Vylder (1976) generalised the homogeneous credibility estimator $\breve{m}_l$ of Hachemeister (1975) to the present model. He defined the homogeneous credibility estimator $\breve{M}$ of $M$ based on $\mathbf{X}$ to be the best estimator of the form $\dot{M} = \mathbf{g}'\mathbf{X}$ that satisfies the unbiasedness constraint $\mathrm{E}\dot{M} = \mathrm{E}M$, that is,

$$\mathbf{g}'\mathrm{E}\mathbf{X} = \mathrm{E}M\,, \tag{4.2}$$

and obtained

$$\breve{M} = \left(\mathrm{Cov}(M, \mathbf{X}')\frac{\mathrm{E}M - \mathrm{Cov}(M, \mathbf{X}')(\mathrm{Cov}\,\mathbf{X})^{-1}\mathrm{E}\mathbf{X}}{\mathrm{E}\mathbf{X}'(\mathrm{Cov}\,\mathbf{X})^{-1}\mathrm{E}\mathbf{X}}\mathrm{E}\mathbf{X}'\right)(\mathrm{Cov}\,\mathbf{X})^{-1}\mathbf{X}\,. \tag{4.3}$$

This result is also given by Dannenburg, Kaas, & Goovaerts (1996).

In this case we have assumed no connection between $\mathrm{E}\mathbf{X}$ and $\mathrm{E}M$. Consequently the best homogeneous linear estimator of $M$ under the constraint (4.2) will have to depend on these mean values. This property of the estimator becomes even more striking when noticing that the fraction in (4.3) does not exist when $\mathrm{E}\mathbf{X} = \mathbf{0}$. In that case the left-hand expression in (4.2) becomes equal to zero, and hence that constraint can be fulfilled only when $\mathrm{E}M = 0$. In that case the inhomogeneous credibility estimator is homogeneous, and hence this estimator is also the homogeneous credibility estimator.

4B. As indicated at the end of Section 2, to avoid that the homogeneous credibility estimator depends on unknown parameters of the means of the estimand and the observations, we have to assume that the mean of the estimand can be expressed as a homogeneous linear function of the means of the observations. Formally such a relation is already implied by (4.2). However, there it is only assumed to hold for one realisation of the values of the means. Consequently, the coefficients will depend on this realisation. However, if it should make sense to use a homogeneous estimator instead of the inhomogeneous credibility estimator, then these means are likely to be unknown, and our main reason for applying a homogeneous estimator is to avoid separate estimation of them. Thus the unbiasedness constraint should hold not only for one realisation of the values of the means, but for a whole set of such realisations.
Now let the vector $\boldsymbol{\tau}$ represent the unknown parameters in $\mathrm{E}\mathbf{X}$ and $\mathrm{E}M$. Then we can express $\mathrm{E}\mathbf{X}$ and $\mathrm{E}M$ as functions of $\boldsymbol{\tau}$,

$$\boldsymbol{\chi}(\boldsymbol{\tau}) = \mathrm{E}\mathbf{X}\,; \qquad \mu(\boldsymbol{\tau}) = \mathrm{E}M\,,$$

and we want

$$\mathbf{g}'\boldsymbol{\chi}(\boldsymbol{\tau}) = \mu(\boldsymbol{\tau}) \tag{4.4}$$

for each possible value of $\boldsymbol{\tau}$. Thus each value of $\boldsymbol{\tau}$ poses a linear constraint on $\mathbf{g}$. To be able to find a $\mathbf{g}$ satisfying all the constraints, we can have at most $n$ independent linear constraints. Let $q$ be the number of independent constraints. Then there must exists an $n \times q$ matrix $\mathbf{Y}$ of full rank $q$, a $q \times 1$ vector $\mathbf{a}$, and a $q \times 1$ vector function $\boldsymbol{\xi}$ of $\boldsymbol{\tau}$ such that

$$\mathrm{E}\mathbf{X} = \mathbf{Y}\boldsymbol{\xi}(\boldsymbol{\tau})\,; \qquad \mathrm{E}M = \mathbf{a}'\boldsymbol{\xi}(\boldsymbol{\tau})\,, \tag{4.5}$$

and thus the constraints (4.4) can be written as

$$(\mathbf{g}'\mathbf{Y} - \mathbf{a}')\boldsymbol{\xi}(\boldsymbol{\tau}) = 0\,. \tag{4.6}$$

As we know that this represents $q$ independent linear constraints, (4.6) is satisfied for all $\boldsymbol{\tau}$ if and only if

$$\mathbf{g}'\mathbf{Y} = \mathbf{a}'\,. \tag{4.7}$$

We define the homogeneous credibility estimator $\ddot{M}$ of $M$ based on **X** to be best homogeneous linear estimator of $M$ based on **X** that satisfies these constraints. The present author finds that this is the only situation where it makes sense to consider homogeneous credibility estimators. This framework includes in particular Hachemeister's regression model.
We see that (4.7) holds if and only if

$$(\mathbf{g}'\mathbf{Y} - \mathbf{a}')\boldsymbol{\beta} = 0$$

for all values of the $q \times 1$ vector $\boldsymbol{\beta}$. We can therefore without loss of generality replace the assumption (4.5) with

$$\mathrm{E}\mathbf{X} = \mathbf{Y}\boldsymbol{\beta}; \qquad \mathrm{E}M = \mathbf{a}'\boldsymbol{\beta},$$

where we do not put any restrictions on the unknown parameter vector $\boldsymbol{\beta}$. For the following we shall make that assumption.

**Theorem 4.2.** *Let the coefficient vector $\gamma$ be defined by*

$$\widetilde{M} = \gamma\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})\boldsymbol{\beta}. \tag{4.8}$$

*Then*

$$\ddot{M} = \gamma'\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})\widehat{\boldsymbol{\beta}}, \tag{4.9}$$

*where $\widehat{\boldsymbol{\beta}}$ is the best linear unbiased estimator of $\boldsymbol{\beta}$ based on* **X**.

**Proof.** Any homogeneous estimator of $M$ that satisfies the constraint (4.7), can be written in the form

$$\dot{M} = \gamma'\mathbf{X} + \mathbf{d}'\mathbf{X},$$

where **d** satisfies the constraint

$$\mathbf{d}'\mathbf{Y} = \mathbf{a}' - \gamma\mathbf{Y}. \tag{4.10}$$

Thus we have to minimise

$$Q = \mathrm{E}(\gamma'\mathbf{X} + \mathbf{d}'\mathbf{X} - M)^2$$

under the constraint (4.10). By application of Theorem 4.1 we obtain

$$\begin{aligned} Q &= \mathrm{Var}(\gamma'\mathbf{X} + \mathbf{d}'\mathbf{X} - M) \\ &= \mathrm{Var}(\widetilde{M} - M + \mathbf{d}'\mathbf{X}) \\ &= \mathrm{Var}(\widetilde{M} - M) + \mathrm{Var}(\mathbf{d}'\mathbf{X}) + 2\,\mathrm{Cov}(\widetilde{M} - M, \mathbf{X}')\mathbf{d} \\ &= \mathrm{Var}(\widetilde{M} - M) + \mathrm{Var}(\mathbf{d}'\mathbf{X})\,. \end{aligned}$$

Thus we have to minimise $\mathrm{Var}(\mathbf{d}'\mathbf{X})$ under the constraint (4.10). But that is the same as finding the best linear unbiased estimator of $(\mathbf{a}' - \gamma'\mathbf{Y})\boldsymbol{\beta}$, and that estimator is $(\mathbf{a}' - \gamma'\mathbf{Y})\widehat{\boldsymbol{\beta}}$. This proves the theorem. Q.E.D.

From regression theory we know that $\widehat{\boldsymbol{\beta}} = (\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y})^{-1}\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{X}$ with $\boldsymbol{\Sigma} = \mathrm{Cov}\,\mathbf{X}$. From Theorem 4.2 we see that we obtain the homogeneous credibility estimator from the inhomogeneous credibility estimator by replacing the constant term by its best linear unbiased estimator. This can be applied to obtain expressions for homogeneous credibility estimators when expressions for the corresponding inhomogeneous credibility estimators are known, e. g. in the Bühlmann-Straub model and Hachemeister's regression model. For the latter model Theorem 4.2 implies that $\widehat{\boldsymbol{\beta}}$ given by (3.5) is the best linear unbiased estimator of $\boldsymbol{\beta}$ based on $\mathbf{X}_1, \dots, \mathbf{X}_k$.
De Vylder (1978) proves a result similar to Theorem 4.2 under Hachemeister's regression model.

4C. Sundt (1987) studied the following credibility regression model, which differs from Hachemeister's model.
We assume that data from $k$ different car models are independent. It is assumed that car model $i$ is characterised by an unknown random risk parameter $\Theta_i$, and that the $\Theta_i$'s are independent and identically distributed.
For car model $i$ we have observed the random variables $X_{i1}, \dots, X_{i,n_i}$. We assume that these variables are conditionally independent given $\Theta_i$, and that

$$\mathrm{E}[\,X_{ij} \mid \Theta_i\,] = m(\Theta_i)\,; \qquad \mathrm{Var}[\,X_{ij} \mid \Theta_i\,] = \frac{v(\Theta_i)}{P_{ij}}$$

$$\lambda = \mathrm{Var}\,m(\Theta_i)\,; \qquad \mathrm{E}m(\Theta_i) = \mathbf{y}_i'\boldsymbol{\beta}\,,$$

where $\mathbf{y}_i$ is a known, non-random $q \times 1$ vector based on technical data on the car model, and $\boldsymbol{\beta}$ is an unknown $q \times 1$ parameter vector. We introduce

$$\varphi = \mathrm{E}v(\Theta_i)\,.$$

We see that if we consider only one car model, then we are within the assumptions of the Bühlmann-Straub model. Thus, as an inhomogeneous credibility estimator does not depend on observations that are independent of the estimand and the other observations (cf. Theorem 6.2 in Sundt (1993)), we obtain from (2.3)–(2.5) that the inhomogeneous credibility estimator $\widetilde{m}_l$, of $m_l(\Theta_l)$ based on the observed $X_{ij}$'s is given by

$$\widetilde{m}_l = \zeta_l X_{l\cdot} + (1-\zeta_l)\mathbf{y}_l'\boldsymbol{\beta} \tag{4.11}$$

with

$$\zeta_l = \frac{P_{l\cdot}}{P_{l\cdot} + \kappa}; \qquad P_{l\cdot} = \sum_{j=1}^{n} P_{lj}; \qquad X_{l\cdot} = \frac{1}{P_{l\cdot}} \sum_{j=1}^{n} P_{lj} X_{lj}; \qquad \kappa = \frac{\varphi}{\lambda},$$

that is, we only have to replace $\mu$ with $\mathbf{y}_l'\boldsymbol{\beta}$ in (2.3).
For the homogeneous credibility estimator $\ddot{m}_l$ we cannot apply the expression from the Bühlmann-Straub model as the car models do not have the same mean. However, from Theorem 4.2 and (4.11) we obtain that

$$\ddot{m}_l = \zeta_l X_{l\cdot} + (1-\zeta_l)\mathbf{y}_l'\widehat{\boldsymbol{\beta}},$$

where $\widehat{\boldsymbol{\beta}}$ is the best linear unbiased estimator of $\boldsymbol{\beta}$ based on the observed $X_{ij}$'s.

## 5 A hierarchical model

5A. For $i = 1, \ldots, k$ let $\mathbf{X}_i$ be an observed random vector related to e.g. an insurance policy, and let $\Theta_i$ be an unknown random parameter representing risk characteristics of that policy. In credibility models like the Bühlmann-Straub model and the regression models of Hachemeister (1975) and Sundt (1987) we have assumed that the $(\mathbf{X}_i, \Theta_i)$'s are independent and the $\Theta_i$'s identically distributed. In a hierarchical extension of such a model we assume that the $(\mathbf{X}_i, \Theta_i)$'s are *conditionally* independent and the $\Theta_i$'s *conditionally* identically distributed given an unknown random *hyper-parameter* H.
We can interpret the policies to be from the same district. The hyperparameter H represents unknown characteristics of that district. We assume that random variables related to different districts are independent and the H's identically distributed.

5B. We shall now discuss a hierarchical extension of the model of Theorem 4.2. Let $\mathbf{X}$ be an observed random $n \times 1$ vector and $M$ an unknown random variable that we want to estimate with the inhomogeneous credibility estimator $\widetilde{M}$ based on $\mathbf{X}$. We also introduce the unknown random hyper-parameter H and assume that

$$\mathrm{E}[\mathbf{X} \mid \mathrm{H}] = \mathbf{Y}\mathbf{b}(\mathrm{H})\,; \qquad \mathrm{E}[\,M \mid \mathrm{H}] = \mathbf{a}'\mathbf{b}(\mathrm{H})\,,$$

where $\mathbf{Y}$ is a non-random $n \times q$ matrix of full rank $q \leq n$, $\mathbf{a}$ is a non-random $q \times 1$ vector, and $\mathbf{b}(\mathrm{H})$ is a random $q \times 1$ vector. We also introduce the structure parameters

$$\begin{aligned} \boldsymbol{\Sigma} &= \mathrm{E}\,\mathrm{Cov}[\mathbf{X} \mid \mathrm{H}\,]\,; \qquad & \boldsymbol{\Psi} &= \mathrm{E}\,\mathrm{Cov}[\,M, \mathbf{X}' \mid \mathrm{H}\,] \\ \boldsymbol{\Xi} &= \mathrm{Cov}\,\mathbf{b}(\mathrm{H})\,; & \boldsymbol{\beta} &= \mathrm{E}\mathbf{b}(\mathrm{H})\,. \end{aligned}$$

The model of Theorem 4.2 appears as a special case with $\boldsymbol{\xi} = \mathbf{0}$. In that case it follows from (4.1) that $\widetilde{M}$ is given by (4.8) with

$$\gamma = (\boldsymbol{\Psi}\boldsymbol{\Sigma}^{-1})'\,.$$

Theorem 6.8 in Sundt (1993) gives a way to generalise this expression for the credibility estimator to the hierarchical case. We simply replace the constant term $(\mathbf{a}' - \gamma'\mathbf{Y})\boldsymbol{\beta}$ with the credibility estimator $\widetilde{S}$ of its generalisation

$$S(\mathrm{H}) = (\mathbf{a}' - \gamma'\mathbf{Y})\mathbf{b}(\mathrm{H})\,,$$

that is, we obtain

$$\widetilde{M} = \gamma'\mathbf{X} + \widetilde{S}\,.$$

For the deduction of an expression for $\widetilde{S}$ we have the same structure of the first and second order moments as in Hachemeister's regression model. Thus we obtain

$$\widetilde{S} = (\mathbf{a}' - \gamma'\mathbf{Y})[\,\boldsymbol{\Delta}\widehat{\mathbf{b}} + (\mathbf{I} - \boldsymbol{\Delta})\boldsymbol{\beta}\,]$$

with

$$\begin{aligned} \widehat{\mathbf{b}} &= (\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y})^{-1}\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{X} \\ \boldsymbol{\Delta} &= \boldsymbol{\Xi}\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y}(\mathbf{I} + \boldsymbol{\Xi}\mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y})^{-1}\,, \end{aligned} \tag{5.1}$$

that is,

$$\widetilde{M} = \gamma'\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})[\,\boldsymbol{\Delta}\widehat{\mathbf{b}} + (\mathbf{I} - \boldsymbol{\Delta})\boldsymbol{\beta}\,]\,. \tag{5.2}$$

As already mentioned, in the special case of the non-hierarchical model we have $\boldsymbol{\Xi} = \mathbf{0}$. In that case we see that $\boldsymbol{\Delta} = \mathbf{0}$, and (5.2) reduces to (4.8) as expected.
In a Bayesian setting, letting $\boldsymbol{\Delta}$ approach $\mathbf{0}$ can be interpreted as obtaining full information about $\mathbf{b}(\mathrm{H})$. The opposite limiting case is when we have no information about $\mathbf{b}(\mathrm{H})$. This is obtained by letting the precision matrix $\boldsymbol{\Pi} = \boldsymbol{\Xi}^{-1}$ approach $\mathbf{0}$. From (5.1) we obtain

$$\boldsymbol{\Delta} = \mathbf{I} - (\boldsymbol{\Pi} + \mathbf{Y}'\boldsymbol{\Sigma}^{-1}\mathbf{Y})^{-1}\boldsymbol{\Pi},$$

and from this we see that $\boldsymbol{\Delta} = \mathbf{I}$ when $\boldsymbol{\Pi} = \mathbf{0}$. In that case (5.2) reduces to

$$\widetilde{M} = \gamma'\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})\widehat{\mathbf{b}},$$

which is in the same form as the homogeneous credibility estimator (4.9).
We have now seen that both the inhomogeneous and the homogeneous credibility estimators of the model of Theorem 4.2 can be obtained as limiting cases of the hierarchical model. The inhomogeneous estimator corresponds to full information about $\mathbf{b}(\mathrm{H})$, and the homogeneous estimator corresponds to no information about $\mathbf{b}(\mathrm{H})$. This was pointed out by Jewell (1975) in the special case of the Bühlmann-Straub model where $P_{ij} = 1$ for all $i, j$.

5C. The best linear unbiased estimator of $\boldsymbol{\beta}$ based on $\mathbf{X}$ is $\widehat{\mathbf{b}}$, and by application of (5.2) and Theorem 4.2 we obtain that the homogeneous credibility estimator of $M$ based on $\mathbf{X}$ is

$$\ddot{M} = \gamma'\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})\widehat{\mathbf{b}}.$$

We now assume that in addition to $\mathbf{X}$ we have observed the collateral data $\mathbf{V}$ independent of $\mathbf{X}$ and H. Let $\mathbf{U} = (\mathbf{X}', \mathbf{V}')'$. By Theorem 6.2 in Sundt (1993) the inhomogeneous credibility estimator of $M$ based on $\mathbf{U}$ is still given by (5.2). However, the homogeneous credibility estimator is now by Theorem 4.2

$$\ddot{M} = \gamma'\mathbf{X} + (\mathbf{a}' - \gamma'\mathbf{Y})[\boldsymbol{\Delta}\widehat{\mathbf{b}} + (\mathbf{I} - \boldsymbol{\Delta})\widehat{\boldsymbol{\beta}}],$$

where $\widehat{\boldsymbol{\beta}}$ denotes the best linear unbiased estimator of $\boldsymbol{\beta}$ based on $\mathbf{U}$.
Gisler (1990) discusses homogeneous credibility estimators within a hierarchical extension of the Bühlmann-Straub model.

**Acknowledgements.**
The author is grateful to Hans Bühlmann for useful feedback, in particular in connection with Theorem 2.1.

## References

Bühlmann, H. (1971). Credibility procedures. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* **1,** 515–525. University of California Press.

Bühlmann, H. & Straub, E. (1970). Glaubwürdigkeit für Schadensätze. *Bulletin of the Association of Swiss Actuaries* **70**, 111–133. English translation by C. E. Brooks (1972): Credibility for loss ratios. *Actuarial Research Clearing House* **1972.2.**

Dannenburg, D. R., Kaas, R., & Goovaerts, M. J. (1996). *Practical actuarial credibility models.* Institute of Actuarial Science and Econometrics, University of Amsterdam

De Vylder, F. (1976). Geometrical credibility. *Scandinavian Actuarial Journal,* 121–149.

De Vylder, F. (1978). Parameter estimation in credibility theory. *ASTIN Bulletin* **10**, 99–112.

Gisler, A. (1990). Credibility theory made easy. *Bulletin of the Swiss Association of Actuaries,* 75–99.

Hachemeister, C. A. (1975). Credibility for regression models with application to trend. In *Credibility: Theory and applications* (ed. P. M. Kahn), pp.129–163. Academic Press, New York.

Jewell, W. S. (1975). The use of collateral data in credibility theory: A hierarchical model. *Giornale dell'Istituto Italiano degli Attuari* **38**, 1–16.

Sundt, B. (1978). On models and methods of credibility. *Statistical Research Report* **7**, Department of Mathematics, University of Oslo.

Sundt, B. (1987). Two credibility regression approaches for the classification of passenger cars in a multiplicative tariff. *ASTIN Bulletin* **17**, 41–70.

Sundt, B. (1993). *An introduction to non-life insurance mathematics.* (3. ed.) Verlag Versicherungswirtschaft e. V., Karlsruhe.

Sundt, B. (1997). Book review. *Bulletin of the Swiss Association of Actuaries,* 89–93.

Taylor, G. C. (1975). Credibility for time-heterogeneous loss ratios. In *Credibility: Theory and applications* (ed. P. M. Kahn), pp. 363–389. Academic Press, New York.

Taylor, G. C. (1977). Abstract credibility. *Scandinavian Actuarial Journal,* 149–168.

Bjørn Sundt
Department of Mathematics
University of Bergen
Johannes Bruns gate 12
N-5008 Bergen
Norway

## Summary

In the present paper we give a historical overview of the development of the theory of homogeneous credibility estimators. We discuss under what conditions it could be interesting to study such estimators, and under these conditions we prove a general result on the relation between homogeneous and inhomogeneous credibility estimators. We finally show how homogeneous and inhomogeneous credibility estimators appear as limiting cases within a hierarchical framework.

## Zusammenfassung

Der vorliegende Artikel liefert einen geschichtlichen Überblick über die Entwicklung der Theorie der homogenen Credibility-Schätzer. Wir gehen der Frage nach, unter welchen Bedingungen eine Untersuchung solcher Schätzer interessant sein könnte; unter diesen Bedingungen erhalten wir ein allgemeines Ergebnis über die Beziehung zwischen homogenen und inhomogenen Credibility-Schätzern. Weiterhin wird aufgezeigt, dass homogene und inhomogene Credibility-Schätzer als Grenzfälle im Rahmen des hierarchischen Modells interpretiert werden können.

## Résumé

Ce travail passe en revue le développement de la théorie des estimateurs de crédibilité hornogènes. Nous discutons quelles sont les conditions sous lesquelles il est intéressant d'étudier de tels estimateurs, et nous démontrons un résultat général de la relation entre les estimateurs de crédibilité homogènes et inhomogènes. Nous montrons comment dans le cadre de la crédibilité hiérarchique les estimateurs de crédibilité homogènes et inhornogènes peuvent être présentés comme des cas limites.