THE PROBLEM OF DETERMINING ESTIMATORS FOR DIFFERENT STRUCTURAL PARAMETERS IN THE CASE OF CREDIBILITY RESULTS FOR WEIGHTED CONTRACTS

Virginia Atanasiu, Bucharest

(Received April 29, 2006)

Abstract. This paper presents and analyzes the estimators of the structural parameters, in the Bühlmann-Straub model, involving complicated mathematical properties of conditional expectations and of conditional covariances. So to enable to use the better linear credibility results obtained in this model, we will provide useful estimators for the structure parameters. From the practical point of view it is stated the attractive property of unbiasedness for these estimators.

Keywords: credibility premium, structure parameters, unbiased estimators, Bühlmann-Straub model

MSC 2000: 62P05

0. Introduction

In this article we first present the Bühlmann-Straub model—see Section 1, which consists of a portfolio of non-life insurance contracts. In Section 1 we will introduce the assumptions of the Bühlmann-Straub model. In that section the optimal linearized credibility premium is derived. It turns out that this procedure does not provide us with a statistic computable from the observations, since the result involves unknown parameters of the structure function. To obtain estimates for these structure parameters for the Bühlmann-Straub model, the contracts are embedded in a collective of identical contracts, all providing independent information on the structure distribution. In Section 2 we provide some useful estimators for the structure parameters. In that section we give unbiased estimators for the structure parameters, such that if the structure parameters in the optimal linearized credibility premium are replaced by these estimators, a homogeneous estimator results. This last estimator can also be shown to be optimal (see Section 3). In Section 3 we show that this last estimator is in fact the optimal linearized homogeneous credibility estimator.

1. The Bühlmann-Straub model

For this model we look upon the portfolio as represented in Diagram 1. We consider a portfolio which can be subdivided in groups consisting of contracts with common risk parameter, as in Diagram 1.

Contracts		1.............. j ....k
Structure variables		θ_j
Observable variables with	p 1	$X_{j1}(w_{j1})$
associated weights	e 2	$X_{j2}(w_{j2})$
	r $\vdots$	$\vdots\ \vdots$
	i $\vdots$	$\vdots\ \vdots$
	o $\vdots$	$\vdots\ \vdots$
	d t	$X_{jt}(w_{jt})$

Diagram 1. Bühlmann-Straub model

Each contract $j = 1, \ldots, k$ is the average of a group of w_{jr} contracts, where w_{jr} is the weight (size) of the group j at time r, with $r = 1, \ldots, t$. So the weight of a "contract" may vary in time, if this weight is equal to the number of proper contracts grouped into an average contract at time r, where $r = 1, \ldots, t$ ($w_{jr} =$ (# of contracts considered having a common risk parameter θ_j), where $r = 1, \ldots, t$ and $j = 1, \ldots, k$). The model consists of the structural variables θ_j and the observable variables X_{jr}, where $j = 1, \ldots, k$ and $r = 1, \ldots, t$. So the contract j consists of the set of variables

$$(\theta_j, \underline{X}'_j) = \theta_j, X_{jr}, r = 1, \ldots, t,$$

where $j = 1, \ldots, k$; the contract indexed by j is a random vector consisting of a random structure parameter θ_j and observations $X_{j1}, X_{j2}, \ldots, X_{jt}$, see Diagram 1:

$$(\theta_j, \underline{X}'_j) = (\theta_j, X_{j1}, \ldots, X_{jt}),$$

where $j = 1, \ldots, k$. Of course the X_{jr} variables represent the average of w_{jr} contracts grouped together at time r, as follows:

$$X_{jr} = \frac{1}{w_{jr}} \sum_{i=1}^{w_{jr}} X_{jr}^{(i)}, \ r = 1, \ldots, t \text{ and } j = 1, \ldots, k.$$

The Bühlmann-Straub assumptions can be formulated as follows:

(BS_1): the contracts $j = 1, \ldots, k$ (the pairs, $(\theta_j, \underline{X}'_j)$ with $j = 1, \ldots, k$) are independent; moreover, for every contract $j = 1, \ldots, k$ and for θ_j fixed, the variables $X_{j1}, \ldots, X_{jt}$ are conditionally independent. The variables $\theta_1, \ldots, \theta_k$ are identically distributed. The observations X_{jr}, $j = 1, \ldots, k$, $r = 1, \ldots, t$ have finite variance.

(BS_2): $E(X_{jr}|\theta_j) = \mu(\theta_j)$, $j = 1, \ldots, k$, $r = 1, \ldots, t$ (we assume that all contracts have common expectation of the claim size as a function $\mu(\cdot)$ of the risk parameter θ_j, where $j = 1, \ldots, k$).

$\mathrm{Var}(X_{jr}|\theta_j) = \sigma^2(\theta_j)/w_{jr}$, $j = 1, \ldots, k$, $r = 1, \ldots, t$, where all $w_{jr} > 0$, with $X_{jr}^{(i)}$, $i = 1, \ldots, w_{jr}$, $j = 1, \ldots, k$, $r = 1, \ldots, t$ satisfying the hypotheses (BS'_1) and (BS'_2):

(BS'_1): for every $j = 1, \ldots, k$ and for θ_j fixed, the variables $X_{jr}^{(i)}$, $i = 1, \ldots, w_{jr}$, $r = 1, \ldots, t$ are conditionally independent and identically distributed. The variables $\theta_1, \ldots, \theta_k$ are identically distributed and the observations $X_{jr}^{(i)}$, $i = 1, \ldots, w_{jr}$, $r = 1, \ldots, t$, $j = 1, \ldots, k$ have finite variance, and

(BS'_2): $E(X_{jr}^{(i)}|\theta_j) = \mu(\theta_j)$, $i = 1, \ldots, w_{jr}$, $r = 1, \ldots, t$, $j = 1, \ldots, k$,

$$\mathrm{Var}(X_{jr}^{(i)}|\theta_j) = \sigma^2(\theta_j),\ i = 1, \ldots, w_{jr}, r = 1, \ldots, t,\ j = 1, \ldots, k.$$

A consequence of the hypothesis (BS_1):

$$\mathrm{Cov}(X_{jr}, X_{jq}|\theta_j) = 0,\ j = \overline{1,k},\ r, q = \overline{1,t},\ r < q.$$

Remarks. 1) $\mu(\theta_j)$ is the pure net risk premium of the contract j, with $j = 1, \ldots, k$.

2) The Bühlmann-Straub assumptions express the common characteristics of the risk under consideration.

3) The weights arise when the contracts are replaced by averages of identical contracts (with the same risk parameter), and the weight then represents the number of such contracts.

The optimal linearized non-homogeneous credibility estimators are given in the following theorem:

Theorem 1.1 (linearized non-homogeneous credibility estimator in the Bühlmann-Straub model)**.** *Under the hypotheses* (BS_1) *and* (BS_2) *of the Bühlmann-Straub model, the following optimal linearized non-homogeneous credibility estimator for* $\mu(\theta_j)$*, for some fixed* j*, is obtained:*

$$M_j^a = \hat{\mu}(\theta_j) = (1 - z_j)m + z_j M_j, \tag{1.1}$$

where $M_j = X_{jw} = \sum_{q=1}^{t} (w_{jq}/w_{j\cdot})X_{jq}$ *denotes the individual estimator for* $\mu(\theta_j)$ *and the resulting credibility factor for contract* j *is given by*

$$z_j = aw_{j\cdot}/(aw_{j\cdot} + s^2)$$

with $a = \mathrm{Var}[\mu(\theta_j)]$, $s^2 = E[\sigma^2(\theta_j)]$, $m = E[\mu(\theta_j)]$ *as usual, where* $w_{j\cdot} = \sum_{q=1}^{t} w_{jq}$, $j = 1, \ldots, k$.

This result can be found in [5]. To be able to use (1.1), one still has to estimate the portfolio characteristics m, s^2, a. Some unbiased estimators are given in the following section.

2. Parameter estimation

Here and in the following (see Section 3) we present the main results leaving the detailed computations to the reader.

The estimators obtained in the previous section contain unknown structure parameters (the credibility premium for the Bühlmann-Straub model involves three unknown parameters: m, s^2 and a). So the expressions for these (pseudo-) estimators are no longer statistics. But since the contracts are embedded in a collective of identical contracts, all providing independent information on the structure distribution, it is possible to give unbiased estimators of these quantities, so we can replace the unknown structure parameters by estimates. In this section, we consider different contracts, each with the same structure parameters m, s^2 and a, so that we can estimate these quantities using the statistics of the different contracts. Some unbiased estimators for the structure parameters m, s^2 and a are given in the following theorem.

Theorem 2.1 (parameter estimation in the Bühlmann-Straub model)**.** *The estimators*

$$\hat{m} = M_0 = X_{zw} = \sum_{j=1}^{k} \frac{z_j}{z\cdot} X_{jw} \quad \left(\text{where } z\cdot = \sum_{j=1}^{k} z_j\right),$$

$$\hat{s}^2 = \frac{1}{k(t-1)} \sum_{j,s} w_{js}(X_{js} - X_{jw})^2,$$

$$\hat{a} = w_{\cdot\cdot}\left[\sum_j w_{j\cdot}(X_{jw} - X_{ww})^2 - (k-1)\hat{s}^2\right] \Big/ \left(w^2 \cdot\cdot - \sum_j w_{j\cdot}^2\right)$$

(where $w_{\cdot\cdot} = \sum\limits_{j=1}^{k} w_{j\cdot} = \sum\limits_{j=1}^{k}\sum\limits_{q=1}^{t} w_{jq}$, $X_{ww} = \sum\limits_{j=1}^{k} \frac{w_{j\cdot}}{w_{\cdot\cdot}} X_{jw}$*) are unbiased estimators of the corresponding structure parameters, i.e.* $E(\hat{m}) = m$, $E(\hat{s}^2) = s^2$, $E(\hat{a}) = a$.

P r o o f. The proof of $E(\hat{m}) = m$ is easy. Using the covariance relations (the relevant covariance relations between the risk premium, the observations and the weighted averages)—see Remark 2.1, we get

$$
\begin{aligned}
k(t-1)E(\hat{s}^2) &= \sum_{j,s} w_{js}[\operatorname{Var}(X_{js}) + \operatorname{Var}(X_{jw}) - 2\operatorname{Cov}(X_{js}, X_{jw})] \\
&= \sum_{j,s} w_{js}\Big[\Big(a + \frac{s^2}{w_{js}}\Big) + \Big(a + \frac{s^2}{w_{j\cdot}}\Big) - 2\Big(a + \frac{s^2}{w_{j\cdot}}\Big)\Big] \\
&= \sum_{j,s} w_{js}\Big[\Big(a + \frac{s^2}{w_{js}}\Big) - \Big(a + \frac{s^2}{w_{j\cdot}}\Big)\Big] \\
&= \Big[\sum_{j,s} w_{js}\Big(\frac{1}{w_{js}} - \frac{1}{w_{j\cdot}}\Big)\Big]s^2 = \Big[\sum_{j,s} w_{js}\frac{1}{w_{js}} - \sum_{j}\frac{1}{w_{j\cdot}}\Big(\sum_{s} w_{js}\Big)\Big]s^2 \\
&= \Big(kt - \sum_{j}\frac{1}{w_{j\cdot}}w_{j\cdot}\Big)s^2 = (kt-k)s^2 = k(t-1)s^2.
\end{aligned}
$$

So $k(t-1)E(\hat{s}^2) = k(t-1)s^2$, that is $E(\hat{s}^2) = s^2$.

The proof of the unbiasedness of $\hat{a}$ is similar. We have

$$
\begin{aligned}
&\Big(w_{\cdot\cdot}^2 - \sum_{j} w_{j\cdot}^2\Big)E(\hat{a}) \\
&= w_{\cdot\cdot}\Big\{\sum_{j} w_{j\cdot}[\operatorname{Var}(X_{jw}) + \operatorname{Var}(X_{ww}) - 2\operatorname{Cov}(X_{jw}, X_{ww})] - (k-1)s^2\Big\} \\
&= w_{\cdot\cdot}\Big\{\sum_{j} w_{j\cdot}\Big[\Big(a + \frac{s^2}{w_{j\cdot}}\Big) + \Big(\frac{s^2}{w_{\cdot\cdot}} + a\sum_{i}\Big(\frac{w_{i\cdot}}{w_{\cdot\cdot}}\Big)^2\Big) - 2\Big(\frac{s^2}{w_{\cdot\cdot}} + a\frac{w_{j\cdot}}{w_{\cdot\cdot}}\Big)\Big] - (k-1)s^2\Big\} \\
&= w_{\cdot\cdot}\Big[a\sum_{j} w_{j\cdot} + s^2\sum_{j} w_{j\cdot}\frac{1}{w_{j\cdot}} + \frac{s^2}{w_{\cdot\cdot}}\sum_{j} w_{j\cdot} + a\frac{1}{w_{\cdot\cdot}^2}\sum_{j} w_{j\cdot}\sum_{i} w_{i\cdot}^2 \\
&\quad - 2\frac{s^2}{w_{\cdot\cdot}}\sum_{j} w_{j\cdot} - 2a\frac{1}{w_{\cdot\cdot}}\sum_{j} w_{j\cdot}^2 - (k-1)s^2\Big] \\
&= w_{\cdot\cdot}\Big[aw_{\cdot\cdot} + ks^2 + \frac{s^2}{w_{\cdot\cdot}}w_{\cdot\cdot} + a\frac{w_{\cdot\cdot}}{w_{\cdot\cdot}^2}\sum_{j} w_{j\cdot}^2 - 2\frac{s^2}{w_{\cdot\cdot}}w_{\cdot\cdot} - 2a\frac{1}{w_{\cdot\cdot}}\sum_{j} w_{j\cdot}^2 - (k-1)s^2\Big] \\
&= aw_{\cdot\cdot}^2 + ks^2w_{\cdot\cdot} + s^2w_{\cdot\cdot} + a\sum_{j} w_{j\cdot}^2 - 2s^2w_{\cdot\cdot} \\
&\quad - 2a\sum_{j} w_{j\cdot}^2 - ks^2w_{\cdot\cdot} + s^2w_{\cdot\cdot} = aw_{\cdot\cdot}^2 - a\sum_{j} w_{j\cdot}^2 = \Big(w_{\cdot\cdot}^2 - \sum_{j} w_{j\cdot}^2\Big)a.
\end{aligned}
$$

So

$$\left(w_{\cdot\cdot}^2 - \sum_j w_{j\cdot}^2\right) E(\hat{a}) = \left(w_{\cdot\cdot}^2 - \sum_j w_{j\cdot}^2\right) a,$$

that is $E(\hat{a}) = a$.

Theorem 2.1 is proved.

R e m a r k 2.1. We start by deriving the relevant covariance relations between the risk premium, the observations and the weighted averages appearing in Theorem 2.1. Under the hypotheses (BS_1)–(BS_2) the following results can be obtained for the conditional expectations and for the covariances:

$$\mathrm{Cov}[\mu(\theta_j), X_{iq}] = \delta_{ij} a, \tag{2.1}$$

$$\mathrm{Cov}(X_{jq}, X_{ir}) = 0 \quad \text{for} \quad j \neq i, \tag{2.2}$$

$$\mathrm{Cov}(X_{jq}, X_{jr}) = a + \delta_{rq} \frac{s^2}{w_{jq}}, \tag{2.3}$$

$$\mathrm{Cov}(X_{jq}, X_{jw}) = \mathrm{Cov}(X_{jw}, X_{jw}) = a + \frac{s^2}{w_{j\cdot}}, \tag{2.4}$$

$$\mathrm{Cov}(X_{jw}, X_{zw}) = \mathrm{Cov}(X_{zw}, X_{zw}) = \frac{a}{z\cdot}, \tag{2.5}$$

$$\mathrm{Cov}(X_{jw}, X_{ww}) = \frac{s^2}{w_{\cdot\cdot}} + a \frac{w_{j\cdot}}{w_{\cdot\cdot}}, \tag{2.6}$$

$$\mathrm{Cov}(X_{ww}, X_{ww}) = \frac{s^2}{w_{\cdot\cdot}} + a \sum_j \left(\frac{w_{j\cdot}}{w_{\cdot\cdot}}\right)^2. \tag{2.7}$$

The proof of these relations: for $i = j$ we have

(2.8)

$$\begin{aligned}
\mathrm{Cov}[\mu(\theta_j), X_{jq}] &= E\{\mathrm{Cov}[\mu(\theta_j), X_{jq}|\theta_j]\} + \mathrm{Cov}\{E[\mu(\theta_j)|\theta_j], E(X_{jq}|\theta_j)\} \\
&= E[\mu(\theta_j)E(X_{jq}|\theta_j)] - \mu(\theta_j)E(X_{jq}|\theta_j)] + \mathrm{Cov}[\mu(\theta_j), \mu(\theta_j)] \\
&= E(0) + \mathrm{Var}[\mu(\theta_j)] = a.
\end{aligned}$$

For $i \neq j$ we have

(2.9)

$$\begin{aligned}
\mathrm{Cov}[\mu(\theta_j), X_{iq}] &= E\{\mathrm{Cov}[\mu(\theta_j), X_{iq}|\theta_j]\} + \mathrm{Cov}\{E[\mu(\theta_j)|\theta_j], E(X_{iq}|\theta_j)\} \\
&= E[\mu(\theta_j)E(X_{iq}|\theta_j) - \mu(\theta_j)E(X_{iq}|\theta_j)] + \mathrm{Cov}[\mu(\theta_j), E(X_{iq})] \\
&= E(0) + \mathrm{Cov}[\mu(\theta_j), m] = 0 + 0 = 0.
\end{aligned}$$

Combining (2.8), (2.9), we obtain (2.1). If $j \neq i$, then we have

(2.10)
$$\begin{aligned}
\operatorname{Cov}(X_{jq}, X_{ir}) &= E[\operatorname{Cov}(X_{jq}, X_{ir}|\theta_j)] + \operatorname{Cov}[E(X_{jq}|\theta_j), E(X_{ir}|\theta_j)] \\
&= E[E(X_{jq}|\theta_j)E(X_{ir}|\theta_j) - E(X_{jq}|\theta_j)E(X_{ir}|\theta_j)] + \operatorname{Cov}[\mu(\theta_j), E(X_{ir})] \\
&= E(0) + \operatorname{Cov}[\mu(\theta_j), m] = 0 + 0 = 0,
\end{aligned}$$

which implies (2.2). Let $r, q = 1, \ldots, t$, $r \neq q$. We have

(2.11)
$$\begin{aligned}
\operatorname{Cov}(X_{jq}, X_{jr}) &= E[\operatorname{Cov}(X_{jq}, X_{jr}|\theta_j)] + \operatorname{Cov}[E(X_{jq}|\theta_j), E(X_{jr}|\theta_j)] \\
&= E[E(X_{jq}|\theta_j) \cdot E(X_{jr}|\theta_j) - E(X_{jq}|\theta_j)E(X_{jr}|\theta_j)] + \operatorname{Cov}[\mu(\theta_j), \mu(\theta_j)] \\
&= E(0) + \operatorname{Var}[\mu(\theta_j)] = 0 + a = a.
\end{aligned}$$

Let $r = q$ $(= 1, \ldots, t)$. We have

$$\begin{aligned}
\text{(2.12)} \quad \operatorname{Cov}(X_{jq}, X_{jq}) &= \operatorname{Var}(X_{jq}) = E[\operatorname{Var}(X_{jq}|\theta_j)] + \operatorname{Var}[E(X_{jq}|\theta_j)] \\
&= E\Big[\frac{\sigma^2(\theta_j)}{w_{jq}}\Big] + \operatorname{Var}[\mu(\theta_j)] = \frac{1}{w_{jq}} s^2 + a = a + \frac{s^2}{w_{jq}}.
\end{aligned}$$

Consequently, from (2.11), (2.12) we get (2.3). According to (2.3) we have

$$\begin{aligned}
\text{(2.13)} \quad \operatorname{Cov}(X_{jq}, X_{jw}) &= \sum_{r=1}^{t} \frac{w_{jr}}{w_{j\cdot}} \operatorname{Cov}(X_{jq}, X_{jr}) = \sum_{r=1}^{t} \frac{w_{jr}}{w_{j\cdot}} \Big(a + \delta_{rq} \frac{s^2}{w_{jq}}\Big) \\
&= \frac{a}{w_{j\cdot}} w_{j\cdot} + \frac{s^2}{w_{j\cdot}} \frac{1}{w_{jq}} \Big(w_{jq} + \sum_{r=1, r\neq q}^{t} \delta_{rq} w_{jr}\Big) \\
&= a + \frac{s^2}{w_{j\cdot}} \frac{1}{w_{jq}} w_{jq} = a + \frac{s^2}{w_{j\cdot}},
\end{aligned}$$

which implies our first assertion. According to (2.13) we have

$$\begin{aligned}
\text{(2.14)} \quad \operatorname{Cov}(X_{jw}, X_{jw}) &= \sum_{q=1}^{t} \frac{w_{jr}}{w_{j\cdot}} \operatorname{Cov}(X_{jq}, X_{jw}) \\
&= \sum_{q=1}^{t} \frac{w_{jq}}{w_{j\cdot}} \Big(a + \frac{s^2}{w_{j\cdot}}\Big) = \frac{a}{w_{j\cdot}} w_{j\cdot} + \frac{s^2}{w_{j\cdot}} \cdot 1 = a + \frac{s^2}{w_{j\cdot}},
\end{aligned}$$

which proves our second assertion. According to (2.13) we have

(2.15)

$$\begin{aligned}
\operatorname{Cov}(X_{jw}, X_{zw}) &= \sum_{q=1}^{t}\sum_{r=1}^{t} \frac{w_{jq}}{w_{j\cdot}}\frac{z_r}{z_\cdot}\operatorname{Cov}(X_{jq}, X_{rw}) \\
&= \sum_{q=1}^{t}\left[\frac{w_{jq}}{w_{j\cdot}}\frac{z_j}{z_\cdot}\operatorname{Cov}(X_{jq}, X_{jw}) + \sum_{r=1, r\neq j}^{t}\frac{w_{jq}}{w_{j\cdot}}\frac{z_r}{z_\cdot}\operatorname{Cov}(X_{jq}, X_{rw})\right] \\
&= \sum_{q=1}^{t}\left[\frac{w_{jq}}{w_{j\cdot}}\frac{z_j}{z_\cdot}\left(a + \frac{s^2}{w_{j\cdot}}\right) + \sum_{r=1, r\neq j}^{t}\frac{w_{jq}}{w_{j\cdot}}\frac{z_r}{z_\cdot}0\right] \\
&= \frac{a}{z_\cdot}\frac{z_j}{w_{j\cdot}}w_{j\cdot}\frac{1}{z_{j\cdot}} = \frac{a}{z_\cdot},
\end{aligned}$$

where

$$\operatorname{Cov}(X_{jq}, X_{rw}) = \sum_{i=1}^{t}\frac{w_{ri}}{w_{r\cdot}}\operatorname{Cov}(X_{jq}, X_{ri}) = \sum_{i=1}^{t}\frac{w_{ri}}{w_{r\cdot}}0 = 0, \text{ if } r \neq j, \tag{2.16}$$

by virtue of the relation (2.2). From (2.15) one obtains our first assertion. According to (2.15) we have

$$\operatorname{Cov}(X_{zw}, X_{zw}) = \sum_{j=1}^{k}\frac{z_j}{z_\cdot}\operatorname{Cov}(X_{jw}, X_{zw}) = \sum_{j=1}^{k}\frac{z_j}{z_\cdot}\frac{a}{z_\cdot} = \frac{a}{z^2_\cdot}z_\cdot = \frac{a}{z_\cdot}. \tag{2.17}$$

From (2.17) one obtains our second assertion. According to (2.4) and (2.16) we have

(2.18)

$$\begin{aligned}
\operatorname{Cov}(X_{jw}, X_{ww}) &= \sum_{q=1}^{t}\sum_{r=1}^{k}\frac{w_{jq}}{w_{j\cdot}}\frac{w_{r\cdot}}{w_{\cdot\cdot}}\operatorname{Cov}(X_{jq}, X_{rw}) \\
&= \sum_{q=1}^{t}\left[\frac{w_{jq}}{w_{j\cdot}}\frac{w_{j\cdot}}{w_{\cdot\cdot}}\operatorname{Cov}(X_{jq}, X_{jw}) + \sum_{r=1, r\neq j}^{k}\frac{w_{jq}}{w_{j\cdot}}\frac{w_{r\cdot}}{w_{\cdot\cdot}}\operatorname{Cov}(X_{jq}, X_{rw})\right] \\
&= \sum_{q=1}^{t}\left[\frac{w_{jq}}{w_{\cdot\cdot}}\left(a + \frac{s^2}{w_{j\cdot}}\right) + \sum_{r=1, r\neq j}^{k}\frac{w_{jq}}{w_{j\cdot}}\frac{w_{r\cdot}}{w_{\cdot\cdot}}0\right] \\
&= a\frac{1}{w_{\cdot\cdot}}w_{j\cdot} + \frac{s^2}{w_{\cdot\cdot}}\frac{w_{j\cdot}}{w_{j\cdot}} = \frac{s^2}{w_{\cdot\cdot}} + a\frac{w_{j\cdot}}{w_{\cdot\cdot}},
\end{aligned}$$

which implies (2.6). Using (2.18), we obtain

$$
\begin{aligned}
(2.19)\qquad \operatorname{Cov}(X_{ww}, X_{ww}) &= \sum_{j=1}^{k} \frac{w_{j\cdot}}{w_{\cdot\cdot}} \operatorname{Cov}(X_{jw}, X_{ww}) \\
&= \sum_{j=1}^{k} \frac{w_{j\cdot}}{w_{\cdot\cdot}} \Big(\frac{s^2}{w_{\cdot\cdot}} + a \frac{w_{j\cdot}}{w_{\cdot\cdot}} \Big) = \frac{s^2}{w_{\cdot\cdot}^2} w_{\cdot\cdot} + a \sum_{j=1}^{k} \Big(\frac{w_{j\cdot}}{w_{\cdot\cdot}} \Big)^2 \\
&= \frac{s^2}{w_{\cdot\cdot}} + a \sum_{j=1}^{k} \Big(\frac{w_{j\cdot}}{w_{\cdot\cdot}} \Big)^2,
\end{aligned}
$$

which gives (2.7).

R e m a r k 2.2. The estimator for a has the weakness that it may assume negative values whereas a is non-negative. Therefore, we replace a by the estimator $a^* = \max(0, \hat{a})$, thus losing unbiasedness, but gaining admissibility. Also note that $\hat{a}$ in Theorem 2.1 might well be negative. Since we want to estimate $\operatorname{Var}[\mu(\theta_j)]$, a more sensible estimator might be $\max(0, \hat{a})$, but this is of course no longer an unbiased estimator.

R e m a r k 2.3. We want to remark that in the case one uses the formula

$$M_j^a = (1 - \hat{z}_j) M_0 + \hat{z}_j M_j$$

one has $E(M_j^a) \neq m$, in the case the estimators from Theorem 2.1 are used, because then $\hat{z}_j$ is dependent of M_j and M_0, $j = 1, \dots, k$. Of course, the attractive property of unbiasedness is lost in this way, but we can still expect the resulting estimators to be good. For instance, when an estimator is a maximum likelihood estimator for a parameter.

R e m a r k 2.4. Theorems 1.1 and 2.1 give the solution to the Bühlmann-Straub model in the case of a non-homogeneous linear estimator for $\mu(\theta_j)$ or, which amounts to the same, for $X_{j,t+1}$, $j = 1, \dots, k$.

R e m a r k 2.5. Note that in the credibility premium for a contract j, the credibility factors z_j also influence the estimator for the overall premium m used. We use X_{zw} rather than X_{ww}, though the latter would be considered more natural by many practicing actuaries. It can be shown that X_{zw} has smaller variance than X_{ww}. In fact X_{ww} has minimal variance in the classical statistical model, but in the credibility model at hand the situation is reversed. To prove that the credibility weighted mean X_{zw}, based on the heterogeneity and the fluctuation of the risk, has minimal mean squared error, we solve

$$\text{(2.20)} \qquad \underset{\underline{\beta}}{\text{Min}}\left\{ \text{Var}\left[\sum_{j=1}^{k} \beta_j X_{jw}\right]\right\} = \underset{\underline{\beta}}{\text{Min}}\left\{ \sum_{j=1}^{k} \beta_j^2 \,\text{Var}(X_{jw})\right\}$$

such that $\sum_{j=1}^{k} \beta_j = 1$ and $\beta_j \geqslant 0$, $j = 1, \ldots, k$, where $\underline{\beta} = (\beta_1, \beta_2, \ldots, \beta_k)'$. Remark that

$$\text{Var}\left[\sum_{j=1}^{k} \beta_j X_{jw}\right] = \sum_{j=1}^{k} \beta_j^2 \,\text{Var}(X_{jw}).$$

Indeed, we have

$$\begin{aligned}\text{Var}\left[\sum_{j=1}^{k} \beta_j X_{jw}\right] &= E\left[\left(\sum_{j=1}^{k} \beta_j X_{jw}\right)^2\right] - E^2\left(\sum_{j=1}^{k} \beta_j X_{jw}\right) \\ &= \sum_{j=1}^{k} \beta_j^2 \,\text{Var}(X_{jw}) + 2 \sum_{1 \leqslant j < j' \leqslant k} \beta_j \beta_{j'} \,\text{Cov}(X_{jw}, X_{j'w}),\end{aligned}$$

where

$$\text{Cov}(X_{jw}, X_{j'w}) = \sum_{q=1}^{t}\sum_{r=1}^{t} \frac{w_{jq}}{w_{j\cdot}}\frac{w_{j'r}}{w_{j'\cdot}} \,\text{Cov}(X_{jq}, X_{j'r}) = \sum_{q=1}^{t}\sum_{r=1}^{t} \frac{w_{jq}}{w_{j\cdot}}\frac{w_{j'r}}{w_{j'\cdot}} 0 = 0,$$

by virtue of the relation (2.2) if $j \neq j'$, and thus we conclude that

$$\text{Var}\left[\sum_{j=1}^{k} \beta_j X_{jw}\right] = \sum_{j=1}^{k} \beta_j^2 \,\text{Var}(X_{jw}).$$

Let j be fixed. Since $\text{Var}(X_{jw}) = \text{Cov}(X_{jw}, X_{jw}) = a + s^2/w_{j\cdot} = a/z_j$ by (2.4), the minimal variance unbiased estimator is found by solving the Lagrange problem

$$\text{(2.21)} \qquad \underset{\alpha,\underline{\beta}}{\text{Min}}\left[\sum_{j=1}^{k} \beta_j^2 \frac{a}{z_j} - 2\alpha\left(\sum_{j=1}^{k} \beta_j - 1\right)\right].$$

The restriction $\sum_{j=1}^{k} \beta_j = 1$ can be written as

$$\text{(2.22)} \qquad \sum_{j=1}^{k} \beta_j - 1 = 0.$$

To deal with the constraint (2.22), we add it to (2.20) with a Lagrange multiplier -2α. Thus the problem (2.21) results. Taking the derivatives with respect to β_j, $j = \overline{1, k}$ leads to the equation

$$2\beta_j \frac{a}{z_j} - 2\alpha = 0, \quad j = 1, \ldots, k.$$

This gives

$$\beta_j = \frac{\alpha z_j}{a}, \quad j = 1, \ldots, k, \tag{2.23}$$

where α still has to be determined in such a way that (2.22) holds, too. Summing all the β_j of (2.23), one gets

$$\frac{\alpha}{a} \sum_{j=1}^{k} z_j = 1,$$

that is,

$$\alpha = \frac{a}{z\cdot},$$

and the resulting value for α, inserted in (2.23), gives

$$\beta_j = \frac{z_j}{z\cdot}, \quad j = 1, \ldots, k.$$

Therefore $z_j/z\cdot$, $j = 1, \ldots, k$ are the optimal weights in the sense that

$$\operatorname*{Min}_{\underline{\beta}} \left(\mathrm{Var}\left[\sum_{j=1}^{k} \beta_j X_{jw} \right] \right) = \mathrm{Var}\left(\sum_{j=1}^{k} \frac{z_j}{z\cdot} X_{jw} \right) = \mathrm{Var}(X_{zw}). \tag{2.24}$$

In view of (2.24) we conclude that

$$\mathrm{Var}(X_{zw}) \leqslant \mathrm{Var}\left(\sum_{j=1}^{k} \beta_j X_{jw} \right)$$

for all $\beta_j \geqslant 0$ with $\sum_{j=1}^{k} \beta_j = 1$. Hence, for $\beta_j = w_{j\cdot}/w_{\cdot\cdot}$, $j = 1, \ldots, k$ we obtain

$$\mathrm{Var}(X_{zw}) \leqslant \mathrm{Var}\left(\sum_{j=1}^{k} \frac{w_{j\cdot}}{w_{\cdot\cdot}} X_{jw} \right) = \mathrm{Var}(X_{ww}).$$

Remark 2.6. One could use another unbiased estimator for the structural parameter a, which really is only a pseudo-estimator, since its definition includes the parameter a to be estimated.

Theorem 2.2 (pseudo-estimator for the heterogeneity parameter)**.** *The estimator*

$$\hat{a} = \frac{1}{k-1}\sum_{j=1}^{k} z_j (M_j - M_0)^2$$

is an unbiased estimator of the heterogeneity parameter a.

P r o o f. Remembering that $M_j = X_{jw}$ and $M_0 = X_{zw}$, so $E(M_j) = E(M_0)$, one gets using the covariance relations (2.4), (2.5):

$$\begin{aligned}
(k-1)E(\hat{a}) &= \sum_j z_j E[(M_j - M_0)^2] \\
&= \sum_j z_j \{E[(M_j - M_0)^2] - [E(M_j) - E(M_0)]^2\} \\
&= \sum_j z_j \{E[(M_j - M_0)^2] - [E(M_j - M_0)]^2\} \\
&= \sum_j z_j \operatorname{Var}(M_j - M_0) = \sum_j z_j \operatorname{Cov}(M_j - M_0, M_j - M_0) \\
&= \sum_j z_j \operatorname{Cov}(X_{jw} - X_{zw}, X_{jw} - X_{zw}) \\
&= \sum_j z_j [\operatorname{Cov}(X_{jw}, X_{jw}) - \operatorname{Cov}(X_{jw}, X_{zw}) - \operatorname{Cov}(X_{zw}, X_{jw}) \\
&\quad + \operatorname{Cov}(X_{zw}, X_{zw})] = \sum_j z_j \left[\left(a + \frac{s^2}{w_{j\cdot}}\right) - \frac{a}{z\cdot} - \frac{a}{z\cdot} + \frac{a}{z\cdot}\right] \\
&= \sum_j z_j \left(a + \frac{s^2}{w_{j\cdot}} - \frac{a}{z\cdot}\right) = \sum_j z_j \left(\frac{a w_{j\cdot} + s^2}{w_{j\cdot}} - \frac{a}{z\cdot}\right) \\
&= \sum_j z_j a \frac{s^2 + a w_{j\cdot}}{a w_{j\cdot}} - \frac{a}{z\cdot}\sum_j z_j = a \sum_j z_j \frac{1}{z_j} - \frac{a}{z\cdot} z\cdot = ak - a = (k-1)a.
\end{aligned}$$

So $(k-1)E(\hat{a}) = (k-1)a$, that is $E(\hat{a}) = a$. Theorem 2.2 is proved.

R e m a r k 2.7. The reason to consider this estimator $\hat{a}$ is that, together with $\hat{s}^2$ as in Theorem 2.1, it provides a nice interpretation of the degree of heterogeneity. It also provides insight into a general procedure of extending these results, to the hierarchical models. First, $\hat{s}^2$ measures the fluctuation of the risk or the heterogeneity s^2 in time, see the definition of s^2. Since $s^2 = E[\sigma^2(\theta_j)]$, the part of the variance describing this fluctuation is measured by the squared differences $(X_{js} - X_{jw})^2$, corrected with their natural weights $w_{js}\colon\ w_{js}(X_{js} - X_{jw})^2$. In total there are k times t results, but k expectations are estimated from the individual data. This gives us an unbiased estimator for the part of the variance describing heterogeneity of the individual risks (see $\hat{s}^2$). Secondly, $\hat{a}$ measures the degree of heterogeneity between the contracts. The square of the difference $(M_j - M_0)^2$ between the individual weighted average result M_j and the collective estimator M_0 (weighted by credibility weights) is the relevant quantity for performing the evaluation of the heterogeneity of the contracts. An unbiased estimator for the variance is then credibility weighted average

$$\hat{a} = \left[\sum_{j=1}^{k} z_j (M_j - M_0)^2\right] / (k-1).$$

The division by $(k-1)$ is due to the fact that we consider k contracts. The overall average is calculated by means of the individual results, so the number of independent terms equals $(k-1)$.

Remark 2.8. In case m in (1.1) is estimated by M_0, we obtain a homogeneous linear combination of all observable variables, giving an unbiased estimate of m. This last estimator can also be shown to be optimal (see Section 3). The next section shows that this happens to give the optimal unbiased homogeneous linearized credibility result.

3. The solution to the Bühlmann-Straub model in the case of homogeneous credibility estimators

Replacing the structure parameter m by an unbiased estimate results in a homogeneous credibility estimator. In Section 3, we will show that this last estimator is in fact the optimal linearized homogeneous credibility estimator. Now, we derive the optimal linearized homogeneous credibility estimator.

Theorem 3.1 (homogeneous credibility estimators in the Bühlmann-Straub model). *The solution to the minimization problem*

$$\underset{c_j}{\text{Min}} E\left\{\left[\mu(\theta_j) - \sum_{j=1}^{k}\sum_{r=1}^{t} c_{jir} X_{ir}\right]^2\right\} \tag{3.1}$$

such that

$$E[\mu(\theta_j)] = \sum_{i,r} c_{jir} E(X_{ir}), \tag{3.2}$$

is

$$M_j^a = (1 - z_j) M_0 + z_j M_j \tag{3.3}$$

with z_j as in Theorem 1.1, where $c_j = (c_{jir})_{i,r}$.

Proof. Let j be fixed. The unbiasedness restriction (3.2) can be written as $\sum_{i,r} c_{jir} = 1$, because $E(X_{ir}) = E[\mu(\theta_j)] = m$.

We insert it in the expectation in (3.1) and add it to the function to be optimized with a Lagrange multiplier $2\alpha/m$. The following problem results:

$$\text{(3.4)} \qquad \underset{c_{j,\alpha}}{\text{Min}}\left(E\left\{\left[\mu(\theta_j) - m - \sum_{i,r} c_{jir}(X_{ir} - m)\right]^2\right\} + 2\alpha\left(1 - \sum_{i,r} c_{jir}\right)\right).$$

Since (3.4) is the minimum of a positive definite quadratic form, it suffices to find a solution with all partial derivatives equal to zero. Taking the derivative with respect to $c_{ji'r'}$ gives for $i' = 1, \ldots, k$, $r' = 1, \ldots, t$:

$$\text{(3.5)} \qquad \alpha + \operatorname{Cov}[\mu(\theta_j), X_{i'r'}] = \sum_{i,r} c_{jir} \operatorname{Cov}(X_{ir}, X_{i'r'}).$$

Using the expressions (2.1), (2.2), (2.3) of these covariances in terms of a and s^2, one obtains the following system of equations:

$$\text{(3.6)} \qquad \alpha + \delta_{i'j} a = \sum_r c_{ji'r}(a + \delta_{rr'} s^2 / w_{i'r}), \quad i' = 1, \ldots, k, \ r' = 1, \ldots, t.$$

Indeed, the right hand side of (3.5) can successively be rewritten as

$$\begin{aligned}
&\sum_r \left[\sum_i c_{jir} \operatorname{Cov}(X_{ir}, X_{i'r'})\right] \\
&\quad = \sum_r \left[c_{ji'r} \operatorname{Cov}(X_{i'r}, X_{i'r'}) + \sum_{i; i \neq i'} c_{jir} \operatorname{Cov}(X_{ir}, X_{i'r'})\right] \\
&\quad = \sum_r \left[c_{ji'r}(a + \delta_{rr'} s^2 / w_{i'r}) + \sum_{i; i \neq i'} c_{jir} 0\right] \\
&\quad = \sum_r c_{ji'r}(a + \delta_{rr'} s^2 / w_{i'r}), \quad i' = 1, \ldots, k, \ r' = 1, \ldots, t.
\end{aligned}$$

These equations can be simplified as follows:

$$\text{(3.7)} \qquad \alpha + \delta_{i'j} a = a c_{ji'\cdot} + s^2 c_{ji'r'} / w_{i'r'}, \ i' = 1, \ldots, k, \ r' = 1, \ldots, t$$

where $c_{ji'\cdot} = \sum_r c_{ji'r}$.

Indeed, the right hand side of (3.6) can be successively rewritten as

$$\begin{aligned}
a c_{ji'\cdot} + s^2 \sum_r \delta_{rr'} c_{ji'r} / w_{i'r} &= a c_{ji'\cdot} + s^2 \left(c_{ji'r'} / w_{i'r'} + \sum_{r; r \neq r'} 0 c_{ji'r} / w_{i'r}\right) \\
&= a c_{ji'\cdot} + s^2 c_{ji'r'} / w_{i'r'}.
\end{aligned}$$

Multiplying each equation by $w_{i'r'}$ and summing these equations over the index r', gives for each i'

$$(\alpha + \delta_{i'j}a)w_{i'.} = c_{ji'.}aw_{i'.} + s^2 c_{ji'.}.$$

So

$$c_{ji'.} = (\alpha + \delta_{i'j}a)w_{i'.}/(s^2 + aw_{i'.}). \tag{3.8}$$

Inserting (3.8) into (3.7) gives an expression for $c_{ji'r'}$:

$$c_{ji'r'} = (\alpha + \delta_{i'j}a)[1 - aw_{i'.}/(aw_{i'.} + s^2)]w_{i'r'}/s^2 = (\alpha + \delta_{i'j}a)(1 - z_{i'})w_{i'r'}/s^2.$$

From this the estimator (3.3) for $\mu(\theta_j)$ becomes

$$\hat{\mu}(\theta_j) = \sum_{i',r'} c_{ji'r'}X_{i'r'} = \sum_{i'r'}[(\alpha + \delta_{i'j}a)(1 - z_{i'})w_{i'r'}/s^2]X_{i'r'},$$

where α still has to be determined in such a way that (3.2) holds, too. Summing all the $c_{ji'.}$ of (3.8), one gets

$$\begin{aligned}
1 &= \sum_{i'}\left(\sum_{r'} c_{ji'r'}\right) = \sum_{i'} c_{ji'.} = \sum_{i'}(\alpha + \delta_{i'j}a)w_{i'.}/(s^2 + aw_{i'.}) \\
&= (\alpha/a)\sum_{i'} aw_{i'.}/(s^2 + aw_{i'.}) + \sum_{i'} \delta_{i'j}z_{i'} \\
&= \alpha\left(\sum_{i'} z_{i'}/a\right) + z_j = \alpha z./a + z_j
\end{aligned}$$

and the resulting value for $\alpha = a(1 - z_j)/z.$, inserted in (3.9), gives after some algebraic manipulations the following optimal estimator for $\mu(\theta_j)$:

$$M_j^a = \hat{\mu}(\theta_j) = (1 - z_j)X_{zw} + z_jX_{jw} \tag{3.9}$$

So the theorem is proved.

Remark 3.1. One likely choice in the minimization problem

$$\underset{g(\cdot)}{\mathrm{Min}} E\{[\mu(\theta_j) - g(X_{j1}, \ldots, X_{jt})]^2\},$$

giving easily computable premiums, is

$$g(X_{j1}, \ldots, X_{jt}) = c_0 + \sum_{i=1}^{k}\sum_{r=1}^{t} c_{jir}X_{ir},$$

leading to the so-called linearized credibility results.

Another possibility is to limit oneself to unbiased homogeneous linear estimators, by requiring additionally $c_0 = 0$ and $E[\mu(\theta_j)] = \sum_{i,r} c_{jir} E(X_{ir})$.

Proceeding in this way one gets homogeneous linear credibility formulae. By the requirement of unbiasedness the sum of the credibility premiums equals the global premium on the top-level.

Remark 3.2. In this section we demonstrated that the estimators obtained for the pure net risk premium on contract level are the best linearized homogeneous credibility estimators for the Bühlmann-Straub model.

Conclusions

This paper completes the solution of the Bühlmann-Straub model in the case of a non-homogeneous linear estimator for $\mu(\theta_j)$, or, which is equivalent, for $X_{j,t+1}$, $j = 1, \ldots, k$.

In view of assumption (BS_1) about the independence of the contracts, it might come as a surprise that the premium for a contract j involves results from other contracts.

A closer look at this assumption reveals that this is so because the other contracts provide additional information on the structure distribution.

For this reason the claim figures of other contracts cannot be ignored when estimating the parameters appearing in the credibility estimate for the contract j.

In this article, the classical Bühlmann model is refined by associating the so-called natural weights with the contracts. These weights arise when the contracts are replaced by averages of identical contracts (with the same risk parameter), and the weight then represents the number of such contracts.

But since the contracts are embedded in a collective of identical contracts, all providing independent information on the structure distribution, we can estimate these structural parameters in the Bühlmann-Straub model using the statistics of the individual contracts.

The above two theorems 1.1 and 2.1 show that it is possible to give unbiased estimators of these quantities (the portfolio characteristics) if we have more than one observation available on the risk parameter.

The article contains a description of the Bühlmann-Straub model, behind a heterogenous portfolio, involving an underlying risk parameter for the individual risks.

Since these risks can now no longer be assumed to be independent, mathematical properties of conditional covariances become useful.

This paper is devoted to the Bühlmann-Straub model allowing the contracts to have different weights (volumes) and the purpose of this article is to get unbiased estimators for the portfolio characteristics.

The mathematical theory provides the means to calculate useful estimators for the structure parameters.

From the practical point of view the property of unbiasedness of these estimators is very appealing and very attractive.

The fact that it is based on complicated mathematics, involving conditional expectations and conditional covariances, need not bother the user more than it does when he applies statistical tools like discriminant analysis, scoring models, SAS and GLIM.

References

[1] *Bühlmann, H.*: Optimale Prämienstufensysteme. Mitt. Verein. Schweiz. Versicherungsmath. *64* (1964), 193–214.

[2] *Bühlmann, H.*: Mathematical Methods in Risk Theory. Springer, Berlin, 1970.

[3] *Daykin, C. D., Pentikäinen, T., Pesonen, M.*: Practical Risk Theory for Actuaries. Chapman & Hall, 1993.

[4] *De Vylder, F.*: Parameter estimation incredibility theory. ASTIN Bulletin *10* (1978), 99–112.

[5] *Goovaerts, M. J., Kaas, R., Van Heerwaarden, A. E., Bauwelinckx, T.*: Effective Actuarial Methods, volume 3. Elsevier Science Publishers (1990), 105–239.

[6] *Norberg, R.*: On optimal parameter estimation in credibility. Insur. Math. Econ. *1* (1982), 73–80.

[7] *Sundt, B.*: On choice of statistics in credibility estimation. Scand. Actuar. J. (1979), 115–123.

[8] *Sundt, B.*: An Introduction to Non-Life Insurance Mathematics. Verlag Versicherungswissenschaft, Karlsruhe, 1984, pp. 22–54.

Author's address: *Virginia Atanasiu*, Academy of Economic Studies, Department of Mathematics, Calea Dorobanţilor nr. 15–17, sector 1 Bucharest, 010552, Romania, e-mail: `virginia_atanasiu@yahoo.com`.