BULETINUL ACADEMIEI DE ŞTIINŢE
A REPUBLICII MOLDOVA. MATEMATICA
Number 3(58), 2008, Pages 18–33
ISSN 1024–7696

# Mathematical models in regression credibility theory

Virginia Atanasiu

**Abstract.** In this paper we give the matrix theory of some regression credibility models and we try to demonstrate what kind of data is needed to apply linear algebra in the regression credibility models. Just like in the case of classical credibility model we will obtain a credibility solution in the form of a linear combination of the individual estimate (based on the data of a particular state) and the collective estimate (based on aggregate USA data). To illustrate the solution with the properties mentioned above, we shall need the well-known representation formula of the inverse for a special class of matrices. To be able to use the better linear credibility results obtained in this study, we will provide useful estimators for the structure parameters, using the matrix theory, the scalar product of two vectors, the norm and the concept of perpendicularity with respect to a positive definite matrix given in advance, an extension of Pythagoras' theorem, properties of the trace for a square matrix, complicated mathematical properties of conditional expectations and of conditional covariances.

**Mathematics subject classification:** 15A03, 15A12, 15A48, 15A52, 15A60, 62P05, 62J12, 62J05.
**Keywords and phrases:** Linearized regression credibility premium, the structural parameters, unbiased estimators.

## Introduction

In this paper we give the matrix theory of some regression credibility models.

The article contains a description of the Hachemeister regression model allowing for effects like inflation.

In Section 1 we give Hachemeister's original model, which involves only one isolated contract. In this section we will give the assumptions of the Hachemeister regression model and the optimal linearized regression credibility premium is derived. Just like in the case of classical credibility model, we will obtain a credibility solution in the form of a linear combination of the individual estimate (based on the data of a particular state) and the collective estimate (based on aggregate USA data). To illustrate the solution with the properties mentioned above, we shall need the well-known representation formula of the inverse for a special class of matrices. 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 Hachemeister's classical model we embed the contract in a collective of contracts, all providing independent information on the structure distribution.

Section 2 describes the classical Hachemeister model. In the classical Hachemeister model, a portfolio of contracts is studied. Just as in Section 1, we will derive the best linearized regression credibility premium for this model and we will provide some useful estimators for the structure parameters, using a well-known representation theorem for a special class of matrices, properties of the trace for a square matrix, the scalar product of two vectors, the norm $||\cdot||_P^2$, the concept of perpendicularity $\perp$ and an extension of Pythagoras' theorem, where $P$ is a positive definite matrix given in advance. So, to be able to use the result from Section 1, one still has to estimate the portfolio characteristics. Some unbiased estimators are given in Section 2. From the practical point of view the attractive property of unbiasedness for these estimators is stated.

## 1 The original regression credibility model of Hachemeister

In the original regression credibility model of Hachemeister, we consider one contract with unknown and fixed risk parameter $\theta$, during a period of $t$ ($\geq 2$) years. The yearly claim amounts are denoted by $X_1, \ldots, X_t$. Suppose $X_1, \ldots, X_t$ are random variables with finite variance. The contract is a random vector consisting of a random structure parameter $\theta$ and observations $X_1, \ldots, X_t$. Therefore, the contract is equal to $(\theta, \underline{X}')$, where $\underline{X}' = (X_1, \ldots, X_t)$. For this model we want to estimate the net premium: $\mu(\theta) = E(X_j|\theta)$, $j = \overline{1,t}$ for a contract with risk parameter $\theta$.

*Remark* 1.1. In the credibility models, the pure net risk premium of the contract with risk parameter $\theta$ is defined as:

$$\mu(\theta) = E(X_j|\theta),\ \forall j = \overline{1,t}. \tag{1.1}$$

Instead of assuming time independence in the pure net risk premium (1.1) one could assume that the conditional expectation of the claims on a contract changes in time, as follows:

$$\mu_j(\theta) = E(X_j|\theta) = \underset{\sim j}{Y}' \underset{\sim}{b}(\theta),\ \forall j = \overline{1,t}, \tag{1.2}$$

where the design vector $\underset{\sim j}{Y}$ is known ($\underset{\sim j}{Y}$ is a column vector of length $q$, the non-random $(q \times 1)$ vector $\underset{\sim j}{Y}$ is known) and where the $\underset{\sim}{b}(\theta)$ are the unknown regression constants ($\underset{\sim}{b}(\theta)$ is a column vector of length $q$).

*Remark* 1.2. Because of inflation we are not willing to assume that $E(X_j|\theta)$ is independent of $j$. Instead we make the regression assumption $E(X_j|\theta) = \underset{\sim j}{Y}' \underset{\sim}{b}(\theta)$.

When estimating the vector $\underset{\sim}{\beta}$ from the initial regression hypothesis $E(X_j) = \underset{\sim j}{Y}' \underset{\sim}{\beta}$ formulated by actuary, Hachemeister found great differences. He then assumed that to each of the states there was related an unknown random risk parameter $\theta$ containing the risk characteristics of that state, and that $\theta$'s from different states were independent and identically distributed. Again considering one particular state, we assume that $E(X_j|\theta) = \underset{\sim j}{Y}' \underset{\sim}{b}(\theta)$, with $E[\underset{\sim}{b}(\theta)] = \underset{\sim}{\beta}$.

*Consequence of the hypothesis* (1.2):

$$\underset{\sim}{\mu}^{(t,1)}(\theta) = E(\underset{\sim}{X}|\theta) = \underset{\sim}{Y}\underset{\sim}{b}(\theta), \tag{1.3}$$

where $\underset{\sim}{Y}$ is a $(t \times q)$ matrix given in advance, the so-called design matrix of full rank $q$ $(q \leq t)$ [the $(t \times q)$ design matrix $\underset{\sim}{Y}$ is known and having full rank $q \leq t$] and where $\underset{\sim}{b}(\theta)$ is an unknown regression vector [$\underset{\sim}{b}(\theta)$ is a column vector of length $q$].

*Observations*. By a suitable choice of the $\underset{\sim}{Y}$ (assumed to be known), time effects on the risk premium can be introduced.

**Examples**. **1**) If the design matrix is for example chosen as follows: $\underset{\sim}{Y} = \underset{\sim}{Y}^{(t,3)} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 2^2 \\ \vdots & \vdots & \vdots \\ 1 & t & t^2 \end{bmatrix}$ we obtain a quadratic inflationary trend: $\mu_j(\theta) = b_1(\theta) + jb_2(\theta) + j^2 b_3(\theta)$, $j = \overline{1,t}$, where $\underset{\sim}{b}(\theta) = (b_1(\theta), b_2(\theta), b_3(\theta))'$. Indeed, by standard computations we obtain: $\underset{\sim}{\mu}^{(t,1)}(\theta) = \underset{\sim}{Y}\underset{\sim}{b}(\theta) = (1b_1(\theta) + 1b_2(\theta) + 1^2 b_3(\theta),\ 1b_1(\theta) + 2b_2(\theta) + 2^2 b_2(\theta), \ldots, 1b_1(\theta) + tb_2(\theta) + t^2 b_3(\theta))'$ and as $\underset{\sim}{\mu}^{(t,1)}(\theta) = (\mu_1(\theta), \mu_2(\theta), \ldots, \mu_t(\theta))'$ results that is established our first assertion.

**2**) If the design matrix is for example chosen as follows: $\underset{\sim}{Y} = \underset{\sim}{Y}^{(t,2)} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & t \end{bmatrix}$ (the last column of 1 is omitted) a linear inflation results: $\mu_j(\theta) = b_1(\theta) + jb_2(\theta)$, $j = \overline{1,t}$, where $\underset{\sim}{b}(\theta) = (b_1(\theta), b_2(\theta))'$. The proof is similar.

After these motivating introductory remarks, we state the model assumptions in more detail.

Let $\underset{\sim}{X} = (X_1, \ldots, X_t)'$ be an observed random $(t \times 1)$ vector and $\theta$ an unknown random risk parameter. We assume that:

$$E(\underset{\sim}{X}|\theta) = \underset{\sim}{Y}\underset{\sim}{b}(\theta). \tag{$H_1$}$$

It is assumed that the matrices:

$$\underset{\sim}{\Lambda} = \mathrm{Cov}[\underset{\sim}{b}(\theta)](\underset{\sim}{\Lambda} = \underset{\sim}{\Lambda}^{(q\times q)}) \tag{$H_2$}$$

$$\utilde{\Phi} = E[\mathrm{Cov}(\utilde{X}|\theta)\ (\utilde{\Phi} = \utilde{\Phi}^{(t\times t)}) \tag{$H_3$}$$

are positive definite. We finally introduce: $E[\utilde{b}(\theta)] = \utilde{\beta}$.

Let $\tilde{\mu}_j$ be the credibility estimator of $\mu_j(\theta)$ based on $\utilde{X}$.

For the development of an expression for $\tilde{\mu}_j$, we shall need the following lemma.

**Lemma 1.1 (Representation formula of the inverse for a special class of matrices)**. *Let $\utilde{A}$ be an $(r\times s)$ matrix and $\utilde{B}$ an $(s\times r)$ matrix. Then*

$$(\utilde{I} + \utilde{A}\utilde{B})^{-1} = \utilde{I} - \utilde{A}(\utilde{I} + \utilde{B}\utilde{A})^{-1}\utilde{B}, \tag{1.4}$$

*if the displayed inverses exist.*

*Proof.* We have

$$\begin{aligned}
\utilde{I} &= \utilde{I} + \utilde{A}\utilde{B} - \utilde{A}\utilde{B} = \utilde{I} + \utilde{A}\utilde{B} - \utilde{A}(\utilde{I} + \utilde{B}\utilde{A})(\utilde{I} + \utilde{B}\utilde{A})^{-1}\utilde{B} = \\
&= (\utilde{I} + \utilde{A}\utilde{B}) - (\utilde{I}\utilde{A} + \utilde{A}\utilde{B}\utilde{A})(\utilde{I} + \utilde{B}\utilde{A})^{-1}\utilde{B} = \\
&= (\utilde{I} + \utilde{A}\utilde{B}) - (\utilde{I} + \utilde{A}\utilde{B})\utilde{A}(\utilde{I} + \utilde{B}\utilde{A})^{-1}\utilde{B}
\end{aligned}$$

giving $\utilde{I} = (\utilde{I} + \utilde{A}\utilde{B})[\utilde{I} - \utilde{A}(\utilde{I} + \utilde{B}\utilde{A})^{-1}\utilde{B}]$ and multiplying this equation from the left by $(\utilde{I} + \utilde{A}\utilde{B})^{-1}$ gives (1.4).

*Observation.* $\utilde{I}$ denotes the $(r\times r)$ identity matrix.

The optimal choice of $\tilde{\mu}_j$ is determined in the following theorem:

**Theorem 1.1.** *The credibility estimator $\tilde{\mu}_j$ is given by:*

$$\tilde{\mu}_j = \utilde{Y}'_j[\utilde{Z}\hat{\utilde{b}} + (\utilde{I} - \utilde{Z})\utilde{\beta}], \tag{1.5}$$

*with:*

$$\hat{\utilde{b}} = (\utilde{Y}'\utilde{\Phi}^{-1}\utilde{Y})^{-1}\utilde{Y}'\utilde{\Phi}^{-1}\utilde{X}, \tag{1.6}$$

$$\utilde{Z} = \utilde{\Lambda}\utilde{Y}'\utilde{\Phi}^{-1}\utilde{Y}(\utilde{I} + \utilde{\Lambda}\utilde{Y}'\utilde{\Phi}^{-1}\utilde{Y})^{-1}, \tag{1.7}$$

*where $\utilde{I}$ denotes the $q\times q$ identity matrix ($\hat{\utilde{b}} = \hat{\utilde{b}}^{(q\times 1)}$; $\utilde{Z} = \utilde{Z}^{(q\times q)}$), for some fixed $j$.*

*Proof.* The credibility estimator $\tilde{\mu}_j$ of $\mu_j(\theta)$ based on $\utilde{X}$ is a linear estimator of the form

$$\tilde{\mu}_j = \gamma_0 + \utilde{\gamma}'\utilde{X}, \tag{1.8}$$

which satisfies the normal equations $\begin{cases} E(\tilde{\mu}_j) = E[\mu_j(\theta)] \\ \mathrm{Cov}(\tilde{\mu}_j, X_j) = \mathrm{Cov}[\mu_j(\theta), X_j] \end{cases}$ where $\gamma_0$ is a scalar constant, and $\underset{\sim}{\gamma}$ is a constant $(t \times 1)$ vector.

The coefficients $\gamma_0$ and $\underset{\sim}{\gamma}$ are chosen such that the normal equations are satisfied.

We write the normal equations as

$$E(\tilde{\mu}_j) = \underset{\sim j}{Y}' \underset{\sim}{\beta}, \tag{1.9}$$

$$\mathrm{Cov}(\tilde{\mu}_j, \underset{\sim}{X}') = \mathrm{Cov}[\mu_j(\theta), \underset{\sim}{X}']. \tag{1.10}$$

After inserting (1.8) in (1.10), one obtains

$$\underset{\sim}{\gamma}' \mathrm{Cov}(\underset{\sim}{X}) = \mathrm{Cov}[\mu_j(\theta), \underset{\sim}{X}'], \tag{1.11}$$

where

$$\begin{aligned} \mathrm{Cov}(\underset{\sim}{X}) &= E[\mathrm{Cov}(\underset{\sim}{X}(\theta)] + \mathrm{Cov}[E(\underset{\sim}{X}(\theta)] = \\ &= \underset{\sim}{\Phi} + \mathrm{Cov}[\underset{\sim}{Y}\,\underset{\sim}{b}(\theta)] = \underset{\sim}{\Phi} + \mathrm{Cov}[\underset{\sim}{Y}\,\underset{\sim}{b}(\theta), (\underset{\sim}{Y}\,\underset{\sim}{b}(\theta))'] = \\ &= \underset{\sim}{\Phi} + \underset{\sim}{Y}\mathrm{Cov}[\underset{\sim}{b}(\theta), (\underset{\sim}{b}(\theta))'\underset{\sim}{Y}'] = \underset{\sim}{\Phi} + \underset{\sim}{Y}\mathrm{Cov}[\underset{\sim}{b}(\theta), (\underset{\sim}{b}(\theta))']\underset{\sim}{Y}' = \\ &= \underset{\sim}{\Phi} + \underset{\sim}{Y}\mathrm{Cov}[\underset{\sim}{b}(\theta)]\underset{\sim}{Y}' = \underset{\sim}{\Phi} + \underset{\sim}{Y}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Y}' \end{aligned}$$

and

$$\begin{aligned} \mathrm{Cov}[\mu_j(\theta), \underset{\sim}{X}'] &= \mathrm{Cov}[\mu_j(\theta), E(\underset{\sim}{X}'(\theta)] = \mathrm{Cov}[\underset{\sim j}{Y}'\underset{\sim}{b}(\theta), (\underset{\sim}{Y}\,\underset{\sim}{b}(\theta))'] = \\ &= \underset{\sim j}{Y}'\mathrm{Cov}[\underset{\sim}{b}(\theta), (\underset{\sim}{b}(\theta))'\underset{\sim}{Y}'] = \underset{\sim j}{Y}'\mathrm{Cov}[\underset{\sim}{b}(\theta), (\underset{\sim}{b}(\theta))']\underset{\sim}{Y}' = \\ &= \underset{\sim j}{Y}'\mathrm{Cov}[\underset{\sim}{b}(\theta)]\underset{\sim}{Y}' = \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}' \end{aligned}$$

and thus (1.11) becomes $\underset{\sim}{\gamma}'(\underset{\sim}{\Phi} + \underset{\sim}{Y}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Y}') = \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'$, from which

$$\begin{aligned} \underset{\sim}{\gamma}' &= \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'(\underset{\sim}{\Phi} + \underset{\sim}{Y}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Y}')^{-1} = \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'[(\underset{\sim}{I} + \underset{\sim}{Y}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1})\underset{\sim}{\Phi}]^{-1} = \\ &= \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}(\underset{\sim}{I} + \underset{\sim}{Y}\,\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1})^{-1}. \end{aligned}$$

Lemma 1.1. now gives

$$\begin{aligned} \underset{\sim}{\gamma}' &= \underset{\sim j}{Y}'\underset{\sim}{\Lambda}\,\underset{\sim}{Y}\,\underset{\sim}{\Lambda}^{-1}[\underset{\sim}{I} - \underset{\sim}{Y}(\underset{\sim}{I} + \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1} \cdot \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}] = \\ &= \underset{\sim j}{Y}'[\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1} - \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}(\underset{\sim}{I} + \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}] = \\ &= \underset{\sim j}{Y}'[\underset{\sim}{I} - \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}(\underset{\sim}{I} + \underset{\sim}{I} \cdot \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1} \cdot \underset{\sim}{I}]\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1} = \\ &= \underset{\sim j}{Y}'(\underset{\sim}{I} + \underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{\Lambda}\,\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1} \end{aligned}$$

and, once more using Lemma 1.1

$$\begin{aligned}\underset{\sim}{\gamma}'\underset{\sim}{X} &= \underset{\sim j}{Y}'(\underset{\sim}{I}+\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{X} = \\ &= \underset{\sim j}{Y}'\underset{\sim}{I}(\underset{\sim}{I}+\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}\underset{\sim}{I})^{-1}\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}\underset{\sim}{\hat{b}} = \underset{\sim j}{Y}'[\underset{\sim}{I}-(\underset{\sim}{I}+\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}\underset{\sim}{I})^{-1}]\underset{\sim}{\hat{b}}\end{aligned}$$

with $\underset{\sim}{\hat{b}}$ given by (1.6). According to Lemma 1.1 we obtain

$$\underset{\sim}{\gamma}'\underset{\sim}{X} = \underset{\sim j}{Y}'\{\underset{\sim}{I}-[\underset{\sim}{I}-\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}(\underset{\sim}{I}+\underset{\sim}{\Lambda}\underset{\sim}{Y}'\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}]\}\cdot\underset{\sim}{\hat{b}} = \underset{\sim j}{Y}'\cdot\underset{\sim}{Z}\cdot\underset{\sim}{\hat{b}},$$

with $\underset{\sim}{Z}$ given by (1.7). Insertion in (1.9) gives

$$\gamma_0 + \underset{\sim j}{Y}'\underset{\sim}{Z}E(\underset{\sim}{\hat{b}}) = \underset{\sim j}{Y}'\underset{\sim}{\beta} \tag{1.12}$$

where

$$\begin{aligned}E(\underset{\sim}{\hat{b}}) &= (\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}E(\underset{\sim}{X}) = (\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}E[E(\underset{\sim}{X}|\theta)] = \\ &= (\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y}E[\underset{\sim}{b}(\theta)] = (\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})^{-1}(\underset{\sim}{Y}\underset{\sim}{\Phi}^{-1}\underset{\sim}{Y})\underset{\sim}{\beta} = \underset{\sim}{\beta}\end{aligned}$$

and thus (1.12) becomes $\gamma_0 + \underset{\sim j}{Y}'\underset{\sim}{Z}\underset{\sim}{\beta} = \underset{\sim j}{Y}'\underset{\sim}{\beta}$ from which $\gamma_0 = \underset{\sim j}{Y}'(\underset{\sim}{I}-\underset{\sim}{Z})\underset{\sim}{\beta}$.

This completes the proof of Theorem 1.1.

## 2 The classical credibility regression model of Hachemeister

In this section we will introduce the classical regression credibility model of Hachemeister, which consists of a portfolio of $k$ contracts, satisfying the constraints of the original Hachemeister model.

The contract indexed $j$ is a random vector consisting of a random structure $\theta_j$ and observations $X_{j1},\dots,X_{jt}$. Therefore the contract indexed $j$ is equal to $(\theta_j, \underline{X}'_j)$, where $\underline{X}'_j = (X_{j1},\dots,X_{jt})$ and $j=\overline{1,k}$ (the variables describing the $j^{th}$ contract are $(\theta_j,\underline{X}'_j)$, $j=\overline{1,k}$). Just as in Section 1, we will derive the best linearized regression credibility estimators for this model.

Instead of assuming time independence in the net risk premium:

$$\mu(\theta_j) = E(X_{jq}|\theta_j), \quad j=\overline{1,k},\ q=\overline{1,t} \tag{2.1}$$

one could assume that the conditional expectation of the claims on a contract changes in time, as follows:

$$\mu_q(\theta_j) = E(X_{jq}|\theta_j) = y_{jq}\beta(\theta_j), \quad j=\overline{1,k},\ q=\overline{1,t}, \tag{2.2}$$

with $y_{jq}$ assumed to be known and $\beta(\cdot)$ assumed to be unknown.

*Observations*: By a suitable choice of the $y_{jq}$, time effects on the risk premium can be introduced.

**Examples**. **1**) If for instance the claim figures are subject to a known inflation $i$, (2.2) becomes:

$$\mu_q(\theta_j) = E(X_{jq}|\theta_j) = (1+i)^q \cdot \beta(\theta_j), \quad j = \overline{1,k},\ q = \overline{1,t}.$$

**2**) If in addition the volume $w_j$ changes from contract to contract, one could introduce the model:

$$\mu_q(\theta_j) = E(X_{jq}|\theta_j) = w_j(1+i)^q \cdot \beta(\theta_j), \quad j = \overline{1,k},\ q = \overline{1,t}$$

where $w_j$ and $i$ are given.

*Consequence of the hypothesis* (2.2):

$$\underline{\mu}^{(t,1)}(\theta_j) = E(\underline{X}_j|\theta_j) = x^{(t,n)}\underline{\beta}^{(n,1)}(\theta_j), \quad = \overline{1,k}, \tag{2.3}$$

where $x^{(t,n)}$ is a matrix given in advance, the so-called design matrix, and where the $\underline{\beta}(\theta_j)$ are the unknown regression constants. Again one assumes that for each contract the risk parameters $\underline{\beta}(\theta_j)$ are the same functions of different realizations of the structure parameter.

*Observations*: By a suitable choice of the $x$, time effects on the risk premium can be introduced.

**Examples**. **1**) If the design matrix is for examples chosen as follows:

$x^{(t,3)} = \begin{bmatrix} 1 & 1 & 1^2 \\ 1 & 2 & 2^2 \\ \vdots & \vdots & \vdots \\ 1 & t & t^2 \end{bmatrix}$, we obtain a quadratic inflationary trend:

$$\mu_q(\theta_j) = \beta_1(\theta_j) + q\beta_2(\theta_j) + q^2\beta_3(\theta_j), \quad j = \overline{1,k},\ q = \overline{1,t}, \tag{2.4}$$

where $\underline{\beta}^{(3,1)}(\theta_j) = (\beta_1(\theta_j), \beta_2(\theta_j), \beta_3(\theta_j))'$, with $j = \overline{1,k}$.

**2**) If the design matrix is for example chosen as follows:

$x^{(t,2)} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ \vdots & \vdots \\ 1 & t \end{bmatrix}$ (the last column of 1) is omitted) a linear inflation results:

$$\mu_q(\theta_j) = \beta_1(\theta_j) + q\beta_2(\theta_j), \quad j = \overline{1,k},\ q = \overline{1,t}, \tag{2.5}$$

where $\underline{\beta}^{(2,1)}(\theta_j) = (\beta_1(\theta_j), \beta_2(\theta_j))'$, with $j = \overline{1,k}$.

For some fixed design matrix $x^{(t,n)}$ of full rank $n$ $(n < t)$, and a fixed weight matrix $v_j^{(t,t)}$, the hypotheses of the Hachemeister model are:

$(H_1)$ The contracts $(\theta_j, \underline{X}'_j)$ are independent, the variables $\theta_1, \ldots, \theta_k$ are independent and identically distributed.

$(H_2)$ $E(\underline{X}_j^{(t,1)}|\theta_j) = x^{(t,n)}\underline{\beta}^{(n,1)}(\theta_j)$, $j = \overline{1,k}$, where $\underline{\beta}$ is an unknown regression vector;

$\mathrm{Cov}(\underline{X}_j^{(t,1)}|\theta_j) = \sigma^2(\theta_j) \cdot v_j^{(t,t)}$, where $\sigma^2(\theta_j) = \mathrm{Var}(X_{jr}|\theta_j)$, $\forall r = \overline{1,t}$ and $v_j = v_j^{(t,t)}$ is a known non-random weight $(t \times t)$ matrix, with $\mathrm{rg} v_j = t$, $j = \overline{1,k}$.

We introduce the structural parameters, which are natural extensions of those in the Bühlmann-Straub model. We have:

$$s^2 = E[\sigma^2(\theta_j)] \tag{2.6}$$

$$a = a^{(n,n)} = \mathrm{Cov}[\underline{\beta}(\theta_j)] \tag{2.7}$$

$$\underline{b} = \underline{b}^{(n,1)} = E[\underline{\beta}(\theta_j)], \tag{2.8}$$

where $j = \overline{1,k}$.

After the credibility result based on these structural parameters is obtained, one has to construct estimates for these parameters. Write: $c_j = c_j^{(t,t)} = \mathrm{Cov}(\underline{X}_j)$, $u_j = u_j^{(n,n)} = (x'v_j^{-1}x)^{-1}$, $z_j = z_j^{(n,n)} = a(a + s^2u_j)^{-1}$ = [the resulting credibility factor for contract $j$], $j = \overline{1,k}$.

Before proving the linearized regression credibility premium, we first give the classical result for the regression vector, namely the GLS-estimator for $\underline{\beta}(\theta_j)$.

**Theorem 2.1 (Classical regression result).** *The vector $\underline{B}_j$ minimizing the weighted distance to the observations $\underline{X}_j$,*

$$d(\underline{B}_j) = (\underline{X}_j - x\underline{B}_j)'v_j^{-1}(\underline{X}_j - x\underline{B}_j),$$

*reads*

$$\underline{B}_j = (x'v_j^{-1}x)^{-1}x'v_j^{-1}\underline{X}_j = u_j x' v_j^{-1}\underline{X}_j\ ,$$

or

$$\underline{B}_j = (x'c_j^{-1}x)^{-1}x'c_j^{-1}\underline{X}_j \quad \textit{in case} \quad c_j = s^2v_j + xax'.$$

*Proof.* The first equality results immediately from the minimization procedure for the quadratic form involved, the second one from Lemma 2.1.

**Lemma 2.1. (Representation theorem for a special class of matrices).** *If $C$ and $V$ are $t \times t$ matrices, $A$ an $n \times n$ matrix and $Y$ a $t \times n$ matrix, and*

$$C = s^2V + YAY',$$

*then*

$$(Y'C^{-1}Y)^{-1} = s^2(Y'V^{-1}Y)^{-1} + A$$

*and*

$$(Y'C^{-1}Y)^{-1}Y'C^{-1} = (Y'V^{-1}Y)^{-1}Y'V^{-1}.$$

We can now derive the regression credibility results for the estimates of the parameters in the linear model. Multiplying this vector of estimates by the design matrix provides us with the credibility estimate for $\underline{\mu}(\theta_j)$, see (2.3).

**Theorem 2.2 (Linearized regression credibility premium)**. *The best linearized estimate of* $E[\underline{\beta}^{(n,1)}(\theta_j)|\underline{X}_j]$ *is given by:*

$$\underline{M}_j = z_j^{(n,n)}\underline{B}_j^{(n,1)} + (I^{(n,n)} - z_j^{(n,n)})\underline{b}^{(n,1)} \tag{2.9}$$

*and the best linearized estimate of* $E[x^{(t,n)}\underline{\beta}^{(n,1)}(\theta_j)|\underline{X}_j]$ *is given by*:

$$x^{(t,n)}\underline{M}_j = x^{(t,n)}[z_j^{(n,n)}\underline{B}_j^{(n,1)} + (I^{(n,n)} - z_j^{(n,n)})\underline{b}^{(n,1)}]. \tag{2.10}$$

*Proof.* The best linearized estimate $\underline{M}_j$ of $E[\underline{\beta}(\theta_j)|\underline{X}_j]$ is determined by solving the following problem

$$\underset{\varepsilon}{Min}\; d(\varepsilon), \tag{2.11}$$

with

$$\begin{aligned} d(\varepsilon) &= ||\underline{\beta}(\theta_j) - (\underline{M}_j + \varepsilon\underline{V})||_p^2 = \\ &= E[(\underline{\beta}(\theta_j) - \underline{M}_j - \varepsilon\underline{V})'P(\underline{\beta}(\theta_j) - \underline{M}_j - \varepsilon\underline{V})], \end{aligned} \tag{2.12}$$

where $\underline{V} = \underline{V}^{(n,1)}$ is a linear combination of 1 and the components of $\underline{X}_j$, $P = P^{(n,n)}$ is a positive definite matrix given in advance and $||\cdot||_p^2$ is a norm defined by: $||\underline{X}||_p^2 = E(\underline{X}'PX)$, with $\underline{X} = \underline{X}^{(n,1)}$ an arbitrary vector.

The theorem holds in case $d'(0) = 0$ for every $\underline{V}$. Standard computations lead to

$$\begin{aligned} d(\varepsilon) &= E[(\underline{\beta}(\theta_j))'P\underline{\beta}(\theta_j)] - E[(\underline{\beta}(\theta_j))'P\underline{M}_j] - \\ &- \varepsilon E[(\underline{\beta}(\theta_j))'P\underline{V}] - E[\underline{M}_j'P\underline{\beta}(\theta_j)] + E[\underline{M}_j'P\underline{M}_j] + \\ &+ \varepsilon E[\underline{M}_j'P\underline{V}] - \varepsilon E[\underline{V}'P\underline{\beta}(\theta_j)] + \varepsilon E[\underline{V}'P\underline{M}_j] + \varepsilon^2 E[\underline{V}'P\underline{V}] \end{aligned} \tag{2.13}$$

The derivative $d'(\varepsilon)$ is given by

$$d'(\varepsilon) = -2E[\underline{V}'P(\underline{\beta}(\theta_j) - \underline{M}_j - \varepsilon\underline{V})] \tag{2.14}$$

Define reduced variables by

$$\underline{\beta}^0(\theta_j) = \underline{\beta}(\theta_j) - E[\underline{\beta}(\theta_j)] = \underline{\beta}(\theta_j) - \underline{b} \tag{2.15}$$

$$\underline{B}_j^0 = \underline{B}_j - E(\underline{B}_j) = \underline{B}_j - \underline{b}, \tag{2.16}$$

$$\underline{X}_j^0 = \underline{X}_j - E(\underline{X}_j) = \underline{X}_j - x\underline{b}. \tag{2.17}$$

Inserting $\underline{M}_j$ from (2.9) in (2.14) for $\varepsilon = 0$, we have to prove that

$$E[\underline{V}'P(\underline{\beta}(\theta_j) - Z_j\underline{B}_j - \underline{b}z_j\underline{b})] = 0, \tag{2.18}$$

for every $\underline{V}$.

Using (2.15) and (2.16), the relation (2.18) can be written as

$$E[\underline{V}'P(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)] = 0, \tag{2.19}$$

for every $\underline{V}$.

But since $\underline{V}$ is an arbitrary vector, with as components linear combinations of 1 and the components of $\underline{X}_j$, it may be written as

$$\underline{V} = \underline{\alpha}_0 + \underline{\alpha}_1^{(n,t)}\underline{X}_j^0. \tag{2.20}$$

Therefore one has to prove that

$$E[(\underline{\alpha}_0' + \underline{X}_j^{0'}\alpha_1')P(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)] = 0, \tag{2.21}$$

for every $\underline{V}$.

Standard computations lead to the following expression for the left hand side

$$\begin{aligned}
&\underline{\alpha}_0'PE[\underline{\beta}^0(\theta_j)] + E[\underline{X}_j^{0'}\alpha_1'P\underline{\beta}^0(\theta_j)] - \underline{\alpha}_0'Pz_jE(\underline{B}_j^0) - E[\underline{X}_j^{0'}\alpha_1'PZ_j\underline{B}_j^0] = \\
&= E[\underline{X}_j^{0'}\alpha_1'P(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)] = E\{Tr[\underline{X}_j^{0'}\alpha_1'P(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)]\} = \\
&= E\{Tr[\alpha_1'P(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)\underline{X}_j^{0'}]\} = Tr\{\alpha_1'PE[(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)\underline{X}_j^{0'}]\},
\end{aligned} \tag{2.22}$$

where we used the fact that $E[\underline{B}^0(\theta_j)] = 0$, $E(\underline{\beta}_j^0) = 0$ and that a scalar random variable trivially equals its trace, and also that $Tr(AB) = Tr(BA)$.

Expression (2.22) is equal to zero, as can be seen by

$$\begin{aligned}
&E[(\underline{\beta}^0(\theta_j) - z_j\underline{B}_j^0)\underline{X}_j^{0'}] = E[\underline{\beta}^0(\theta_j)\underline{X}_j^{0'}] - z_jE(\underline{B}_j^0\underline{X}_j^{0'}) = \\
&= \text{Cov}\,[\underline{\beta}^0(\theta_j), \underline{X}_j^{0'}] - z_j\text{Cov}(\underline{B}_j^0, \underline{X}_j^{0'}) = \\
&= \text{Cov}[\underline{\beta}(\theta_j), \underline{X}_j] - z_j\text{Cov}(\underline{B}_j, \underline{X}_j) = \\
&= ax' - z_j(a + s^2u_j)x' = ax' - a(a + s^2u_j)^{-1}(a + s^2u_j)x' = \\
&= ax' - ax' = 0.
\end{aligned} \tag{2.23}$$

This proves (2.9), (2.10) follows by replacing $P$ in (2.12) by $x'Px$. So repeating the same reasoning as above we arrive at (2.10).

*Remark* 2.1. Here and in the following we present the main results leaving the detailed computations to the reader.

*Remark* 2.2. From (2.9) we see that the credibility estimates for the parameters of the linear model are given as the matrix version of a convex mixture of the classical regression result $\underline{B}_j$ and the collective result $\underline{b}$.

Theorem 2.2 concerns a special contract $j$. By the assumptions, the structural parameters $a, \underline{b}$ and $s^2$ do not depend on $j$. So if there are more contracts, these parameters can be estimated.

Every vector $\underline{B}_j$ gives an unbiased estimator of $\underline{b}$. Consequently, so does every linear combination of the type $\Sigma\alpha_j\underline{B}_j$, where the vector of matrices $(\alpha_j^{(n,n)})_{j=\overline{1,k}}$, is such that:

$$\sum_{j=1}^{k}\alpha_j^{(n,n)} = I^{(n,n)}. \tag{2.24}$$

The optimal choice of $\alpha_j^{(n,n)}$ is determined in the following theorem:

**Theorem 2.3 (Estimation of the parameters $\underline{b}$ in the regression credibility model)**. *The optimal solution to the problem*

$$\underset{\underline{\alpha}}{Min}\, d(\alpha), \tag{2.25}$$

*where:*

$$d(\underline{\alpha}) = \left\|\underline{b} - \sum_j \alpha_j \underline{B}_j\right\|_p^2 \underline{\underline{def}}\; E\left[\left(\underline{b} - \sum_j \alpha_j\underline{B}_j\right)' P\left(\underline{b} - \sum_j \alpha_j \underline{B}_j\right)\right]$$

*(the distance from* $\left(\sum_j \alpha_j\underline{B}_j\right)$ *to the parameters* $\underline{b}$*),* $P = P^{(n,n)}$ *a given positive definite matrix (P is a non-negative definite matrix), with the vector of matrices* $\underline{\alpha} = (\alpha_j)_{j=\overline{1,k}}$ *satisfying (2.24), is:*

$$\hat{\underline{b}}^{(n,1)} = Z^{-1}\sum_{j=1}^{k} z_j \underline{B}_j, \tag{2.26}$$

*where* $Z = \sum_{j=1}^{k} z_j$ *and* $z_j$ *is defined as:* $z_j = a(a + s^2u_j)^{-1}$, $j = \overline{1,k}$.

*Proof.* Using the norm $||X||_p^2 = E(X'PX)$ and the perpendicularity concept $\perp$ of two vectors $\underline{X}^{(n,1)}$ and $\underline{Y}^{(n,1)}$ defined by $\underline{X} \perp \underline{Y}$ iff $E(\underline{X}'P\underline{Y}) = 0$, we see that it is sufficient to prove that for all feasible $\underline{\alpha}$

$$(\hat{\underline{b}} - \underline{b}) \perp \left(\sum_j \alpha_j \underline{B}_j - \hat{\underline{b}}\right), \tag{2.27}$$

since then according to an extension of Pythagoras$'$ theorem

$$\underline{X} \perp \underline{Y} \Leftrightarrow ||X+Y||_p^2 = ||X||_p^2 + ||Y||_p^2,$$

we have

$$\begin{aligned} ||\underline{b} - \sum_j \alpha_j \underline{B}_j||_p^2 &= ||\underline{b} - \hat{\underline{b}} + \hat{\underline{b}} - \sum_j \alpha_j \underline{B}_j||_p^2 = \\ &= ||\underline{b} - \hat{\underline{b}}||_p^2 + ||\hat{\underline{b}} - \sum_j \alpha_j \underline{\beta}_j||_p^2 \end{aligned} \tag{2.28}$$

so for every choice of $\underline{\alpha}$ one gets

$$||\underline{b}-\hat{\underline{b}}||_p^2 \leq ||\underline{b}-\sum_j \alpha_j \underline{B}_j||_p^2. \tag{2.29}$$

So let us show now that (2.27) holds. It is clear that

$$\hat{\underline{b}}-\sum_j \alpha_j \underline{B}_j = Z^{-1}\sum_j z_j \underline{B}_j - \sum \alpha_j \underline{B}_j = \sum_j (Z^{-1}Z_j-\alpha_j)\underline{B}_j = \sum_j \gamma_j \underline{B}_j \tag{2.30}$$

with

$$\sum_j \gamma_j = \sum_j (Z^{-1}z_j-\alpha_j) = Z^{-1}\sum_j z_j - \sum_j \alpha_j = Z^{-1}Z - I = I - I = 0, \tag{2.31}$$

where $\gamma_j = Z^{-1}z_j-\alpha_j$, $j=\overline{1,k}$. To prove (2.27), we have to show that

$$\left(\sum_j \gamma_j \underline{B}_j\right) \perp (\hat{\underline{b}}-\underline{b}) \tag{2.32}$$

so that

$$E\left[\left(\sum_j \gamma_j \underline{B}_j\right)' P(\hat{\underline{b}}-\underline{b})\right] = 0. \tag{2.33}$$

The left hand side of (2.33) can successively be rewritten as follows

$$\begin{aligned}
&E\left[\left(\sum_j \underline{B}_j'\gamma_j'\right) P(\hat{\underline{b}}-\underline{b})\right] = \sum_j E(\underline{B}_j'\gamma_j'P\hat{\underline{b}}^0) = \\
&= \sum_j [E(\underline{B}_j'\gamma_j'P\hat{\underline{b}}^0) - \underline{b}'\gamma_j'PE(\hat{\underline{b}}^0)] = \\
&= \sum_j [E(\underline{B}_j'\gamma_j'P\hat{\underline{b}}^0)) - E(\underline{b}'\gamma_j'P\hat{\underline{b}}^0)] = \sum_j E[(\underline{B}_j'-\underline{b}')\gamma_j'P\hat{\underline{b}}^0] = \\
&= \sum_j E[(\underline{B}_j'-E(\underline{B}_j'))\gamma_j'P\hat{\underline{b}}^0] = \sum_j E(\underline{B}_j'^0\gamma_j'P\hat{\underline{b}}^0) = \\
&= \sum_j E(\underline{B}_j'^0\gamma_j'PZ^{-1}\cdot \sum_i z_i\underline{B}_i^0) = \sum_{j,i} E(\underline{B}_j'^0\gamma_j'PZ^{-1}z_i\underline{B}_i^0) = \\
&= \sum_{j,i} E[Tr(\underline{B}_j'^0\gamma_j'PZ^{-1}z_i\underline{B}_i^0)] = \sum_{j,i} E[Tr(\gamma_j'PZ^{-1}z_i\underline{B}_i^0\underline{B}_j'^0)] = \\
&= \sum_{j,i} Tr[\gamma_j'PZ^{-1}z_iE(\underline{B}_i^0\underline{B}_j'^0)] = \sum_{j,i} Tr[\gamma_j'PZ^{-1}z_i\mathrm{Cov}\,(\underline{B}_i^0,\underline{B}_j^0)] = \\
&= \sum_{j,i} Tr[\gamma_j'PZ^{-1}z_i\mathrm{Cov}\,(\underline{B}_i\underline{B}_j)] = \sum_{j,i} Tr[\gamma_j'PZ^{-1}z_i\delta_{ij}(a+s^2u_j)] =
\end{aligned}$$

$$\begin{aligned}
&= \sum_j Tr[\gamma_j' P Z^{-1} z_j (a + s^2 u_j)] = \\
&= \sum_j Tr[\gamma_j' P Z^{-1} z_j a (a + s^2 u_j)^{-1} (a + s^2 u_j)] = \\
&= \sum_j Tr(\gamma_j' P Z^{-1} a) = Tr\left[ (\sum_j \gamma_j') P Z^{-1} a \right] = \\
&= Tr(0 P Z^{-1} a) = Tr(0) = 0,
\end{aligned} \tag{2.34}$$

where $\hat{\underline{b}^0} = \hat{\underline{b}} - E(\hat{\underline{b}}) = \hat{\underline{b}} - \underline{b}$, $\underline{B}_j'^0 = \underline{B}_j' - E(\underline{B}_j') = \underline{B}_j' - \underline{b}'$ are the reduced variables. In (2.34) we used the fact that $E(\hat{\underline{b}^0}) = 0$ and that a scalar random variable trivially equals its trace, and also that $Tr(AB) = Tr(BA)$. The proof is complete.

**Theorem 2.4 (Unbiased estimator for $s^2$ for each contract group).** *In case the number of observations $t_j$ in the $j^{th}$ contract is larger than the number of regression constants $n$, the following is an unbiased estimator of $s^2$:*

$$\hat{s}_j^2 = \frac{1}{t_j - n} (\underline{X}_j - x_j \underline{B}_j)' (\underline{X}_j - x_j \underline{B}_j). \tag{2.35}$$

**Corollary (Unbiased estimator for $s^2$ in the regression model).** *Let $K$ denote the number of contracts $j$, with $t_j > n$. The $E(\hat{s}^2) = s^2$, if:*

$$\hat{s}^2 = \frac{1}{K} \sum_{j; t_j > n} \hat{s}_j^2. \tag{2.36}$$

For $a$, we give an unbiased pseudo-estimator, defined in terms of itself, so it can only be computed iteratively:

**Theorem 2.5 (Pseudo-estimator for a).** *The following random variable has expected value $a$:*

$$\hat{a} = \frac{1}{k-1} \sum_j z_j (\underline{B}_j - \hat{\underline{b}})(\underline{B}_j - \hat{\underline{b}})'. \tag{2.37}$$

*Proof.* By standard computations we obtain

$$E(\hat{a}) = \frac{1}{k-1} \sum_j z_j [E(\underline{B}_j \underline{B}_j') - E(\underline{B}_j \hat{\underline{b}}') - E(\hat{\underline{b}}\, \underline{B}_j') + E(\hat{\underline{b}}\, \hat{\underline{b}}')]. \tag{2.38}$$

Since

$$\begin{aligned}
&\mathrm{Cov}(\underline{B}_j) = \mathrm{Cov}(\underline{B}_j, \underline{B}_j') = \mathrm{Cov}[u_j x' v_j^{-1} \underline{X}_j, (u_j x' v_j^{-1} \underline{X}_j)'] = \\
&= u_j x' v_j^{-1} \mathrm{Cov}(\underline{X}_j) v_j^{-1} x u_j' = u_j x' v_j^{-1} (s^2 v_j + x a x') v_j^{-1} x u_j' = a + s^2 u_j,
\end{aligned}$$

results that

$$E(\underline{B}_j\, \underline{B}_j') = \mathrm{Cov}(\underline{B}_j) + E(\underline{B}_j) E(\underline{B}_j') = a + s^2 u_j + \underline{b}\, \underline{b}', \tag{2.39}$$

where $E(\underline{B}_j) = E[E(\underline{B}_j|\theta_j)] = E(\underline{\beta}(\theta_j)] = \underline{b}$. Since

$$\begin{aligned}\mathrm{Cov}(\underline{B}_j, \hat{\underline{b}}') &= \mathrm{Cov}(\underline{B}_j, Z^{-1}\sum_i z_i\underline{B}_i) = \left(\mathrm{Cov}\left(\sum_i Z^{-1}z_i\underline{B}_i, \underline{B}_j\right)\right)' = \\ &= \sum_i \mathrm{Cov}(\underline{B}_j, \underline{B}_i)z_i' \cdot (Z')^{-1} = \sum_i \delta_{ij}(a+s^2u_j)z_i'(Z')^{-1} = (a+s^2u_j)z_j'(Z')^{-1},\end{aligned}$$

results that

$$E(\underline{B}_j\hat{\underline{b}}') = \mathrm{Cov}(\underline{B}_j, \hat{\underline{b}}') + E(\underline{B}_j)E(\hat{\underline{b}}') = (a+s^2u_j)z_j'(Z')^{-1} + \underline{b}\,\underline{b}', \tag{2.40}$$

where

$$E(\hat{\underline{b}}) = E\left(Z^{-1}\sum_j z_j\underline{B}_j\right) = Z^{-1}\left(\sum_j z_j\right)E(\underline{B}_j) = Z^{-1}Z\underline{b} = \underline{b}.$$

Since

$$\begin{aligned}&\mathrm{Cov}(\hat{\underline{b}}, \underline{B}_j) = (\mathrm{Cov}(\underline{B}_j, \hat{\underline{b}}))' = [(a+s^2u_j)z_j'(Z')^{-1}]' = \\ &= Z^{-1}z_j(a+s^2u_j) = Z^{-1}a(a+s^2u_j)^{-1}(a+s^2u_j) = Z^{-1}a,\end{aligned}$$

results that

$$E(\hat{\underline{b}}\underline{B}_j') = \mathrm{Cov}(\hat{\underline{b}}, \underline{B}_j) + E(\hat{\underline{b}})E(\underline{B}_j') = Z^{-1}a + \underline{b}\,\underline{b}' \tag{2.41}$$

Since

$$\begin{aligned}\mathrm{Cov}(\hat{\underline{b}}) &= \mathrm{Cov}(\hat{\underline{b}}, \hat{\underline{b}}) = \mathrm{Cov}(Z^{-1}\sum_i z_i\underline{B}_i, Z^{-1}\sum_j z_j\underline{B}_j) = \\ &= Z^{-1}\sum_i z_i\left(\sum_j \mathrm{Cov}\,(\underline{B}_j, \underline{B}_i)z_j'\right)\cdot(Z')^{-1} = \\ &= Z^{-1}\sum_i z_i\left(\sum_j \delta_{ij}(a+s^2u_j)z_j'\right)(Z')^{-1} = \\ &= Z^{-1}\left(\sum_i z_i(a+s^2u_i)z_i'\right)(Z')^{-1} = \\ &= Z^{-1}\left(\sum_i a(a+s^2u_i)^{-1}(a+s^2u_i)z_i'\right)(Z')^{-1} = Z^{-1}aZ'(Z')^{-1} = Z^{-1}a\end{aligned}$$

results that

$$E(\hat{\underline{b}}\,\hat{\underline{b}}') = \mathrm{Cov}(\hat{\underline{b}}) + E(\hat{\underline{b}})E(\hat{\underline{b}}') = Z^{-1}a + \underline{b}\,\underline{b}'\ . \tag{2.42}$$

Now (2.37) follows from (2.38), (2.39), (2.40), (2.41) and (2.42).

*Remark* 2.3. Another unbiased estimator for $a$ is the following:

$$\hat{a} = \frac{1}{(w^2. - \sum w_j^2)} \left\{ \frac{1}{2} \sum_{i,j} w_i w_j (\underline{B}_i - \underline{B}_j)(\underline{B}_i - \underline{B}_j)' - \hat{s}^2 \sum_{j=1}^{k} w_j (w. - w_j) u_j \right\}, \tag{2.43}$$

where $w_j$ is the volume of the risk for the $j^{th}$ contract, $j = \overline{1,k}$ and $w. = \sum_j w_j$.

*Proof.* Complicate and tedious computations lead to

$$(w^2. - \sum_j w_j^2) E(\hat{a}) = \frac{1}{2} \left\{ \sum_{i,j} w_i w_j E[(\underline{B}_i - \underline{B}_j) \cdot (\underline{B}_i - \underline{B}_j)'] \right\} - E(\hat{s}^2) \cdot$$

$$\cdot \sum_j w_j (w. - w_j) u_j = \frac{1}{2} \{ \sum_{i,j} w_i w_j [E(\underline{B}_i\ \underline{B}_i') - E(\underline{B}_i\ \underline{B}_j') -$$

$$-E(\underline{B}_j\ \underline{B}_i') + E(\underline{B}_j\ \underline{B}_j')]\} - s^2 \left( \sum_j w_j w. u_j - \sum_j w_j^2 u_j \right) =$$

$$= \tfrac{1}{2} \{ \sum_{i,j} w_i w_j [a + s^2 u_i + \underline{b}\, \underline{b}' - \delta_{ij}(a + s^2 u_j) - \underline{b}\, \underline{b}' - \delta_{ij}(a + s^2 u_j) -$$

$$-\underline{b}\, \underline{b}' + a + s^2 u_j + \underline{b}\, \underline{b}']\} - s^2 \sum_j w_j w. u_j + s^2 \sum_j w_j^2 u_j =$$

$$= \tfrac{1}{2} \cdot 2w.w.a + \tfrac{1}{2} s^2 \sum_i w_i u_i w. - \frac{1}{2} 2 \sum w_j \cdot \sum_i w_i \delta_{ij} (a + s^2 u_j) +$$

$$+ \tfrac{1}{2} s^2 \sum_j w_j u_j w. - s^2 \sum w_j u_j w. + s^2 \sum_j w_j^2 u_j =$$

$$= w^2.a - \sum_j w_j^2 a = (w^2. - \sum_j w_j^2) a$$

Thus we have proved our assertion.

*Observation.* This estimator is a statistic; it is not a pseudo-estimator. Still, the reason to prefer (2.37) is that this estimator can easily be generalized to multi-level hierarchical models. In any case, the unbiasedness of the credibility premium disappears even if one takes (2.43) to estimate $a$.

## 3 Conclusions

The article contains a credibility solution in the form of a linear combination of the individual estimate (based on the data of a particular state) and the collective estimate (based on aggregate USA data). This idea is worked out in regression credibility theory.

In case there is an increase (for instance by inflation) of the results on a portfolio, the risk premium could be considered to be a linear function in time of the type $\beta_0(\theta) + t\beta_1(\theta)$. Then two parameters $\beta_0(\theta)$ and $\beta_1(\theta)$ must be estimated from the observed variables. This kind of problem is named regression credibility. This model arises in cases where the risk premium depends on time, e.g. by inflation. The one could assume a linear effect on the risk premium as an approximation to the real growth, as is also the case in time series analysis.

These regression models can be generalized to get credibility models for general regression models, where the risk is characterized by outcomes of other related variables.

This paper contains a description of the Hachemeister regression model allowing for effects like inflation. If there is an effect of inflation, it is contained in the claim figures, so one should use estimates based on these figures instead of external data. This can be done using Hachemeister's regression model.

In this article the regression credibility result for the estimates of the parameters in the linear model is derived. After the credibility result based on the structural parameters is obtained, one has to construct estimates for these parameters.

The matrix theory provided the means to calculate useful estimators for the structure parameters. The property of unbiasedness of these estimators is very appealing and very attractive from the practical point of view.

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

## References

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

[2] Goovaerts M.J., Kaas R., Van Heerwaarden A.E., Bauwelinckx T. *Effective Actuarial Methods.* Elsevier Science Publishers B.V., **3** 1990.

[3] Hachemeister C.A. *Credibility for regression models with application to trend*; in *Credibility, theory and application.* Proceedings of the Berkeley Actuarial Research Conference on credibility; Academic Press, New York, 1975, 129–163.

[4] Sundt B. *On choice of statistics in credibility estimation*, Scandinavian Actuarial Journal, 1979, 115–123

[5] Sundt B. *An Introduction to Non-Life*, Insurance Mathematics, volume of the "Mannheim Series", 1984, 22–54.

[6] Sundt B. *Two credibility regression approaches for the classification of passenger cars in a multiplicative tariff*, ASTIN Bulletin, **17**, 1987b, 41–69

Academy of Economic Studies
Department of Mathematics
Calea Dorobanţilor nr. 15-17
Sector 1 Bucharest, 010552
Romania
E-mail: *virginia_atanasiu@yahoo.com*

*Received September 27, 2006*