ESTIMATION OF THE STRUCTURAL PARAMETERS IN THE GENERAL SEMI-LINEAR CREDIBILITY MODEL

Virginia ATANASIU[1]

O lucrare originală care prezintă şi analizează estimatorii parametrilor structurali din modelul de credibilitate semi-liniară, implicând proprietăţi matematice complicate ale valorilor medii condiţionate şi ale covarianţelor condiţionte.

Deci pentru a putea folosi rezultatele superioare de credibilitate semi-liniară, obţinute in acest model, vom oferi estimatori utili ai parametrilor de structură.

Din punct de vedere practic, este evidenţiată proprietatea atractivă de nedeplasare a acestor estimatori.

An original paper which presents and analyses the estimators of the structural parameters, in the semi-linear credibility model involving complicated mathematical properties of conditional expectations and of conditional covariances.

Thus, to be able to use the superior semi-linear credibility results obtained in this model, we will provide useful estimators for the structure parameters.

From the practical point of view, the attractive property of the unbiasedness of these estimators is highlighted.

Key words: contracts, unbiased estimators, structure parameters, semi-linear credibility theory.

AMS Subject Classification: 62P05.

1. Introduction

In this article we first give the general semi-linear credibility model (see **Section 1**), which involves only one isolated contract.

We derive the optimal linearized credibility estimate for the risk premium for this case and we consider **as applications** of this result: **1**) the special semi-linear credibility model (obtained from the general semi-linear credibility model for n = 1), **2**) the approximation to $\mu_0(\theta)$-the net premium for a contract with risk parameter θ-based on a unique optimal approximating function f, **3**) the special hierarchical semi-linear credibility model.

[1] Lecturer , Mathematics Department, Academy of Economic Studies, Bucharest, ROMANIA

It turns out that this procedure does not provide us with a statistics computable from the observations, since the result involves unknown parameters of the structure function.

To obtain estimates for these structure parameters, for the general semi-linear credibility model we embed the contract in a collective of contracts, all providing independent information on the structure distribution (see **Section 2**). We first derive the unbiased estimators for the structure parameters of the special semi-linear credibility model. We close this section giving as **application** of this estimation the unbiased estimators obtained for the structural parameters of the general semi-linear credibility model.

Section 2
The semi-linear credibility model

Consider a finite sequence $\theta, X_1, ..., X_t, X_{t+1}$ of random variables. Assume that for fixed θ, the variables $X_1, ..., X_{t+1}$ are conditionally independent and identically distributed (conditionally i.i.d). The variables $X_1, .., X_t$ are observable and θ is the structure variable. The variable X_{t+1} is considered as being not (yet) observable. We assume that $f_p(X_r), p = \overline{0,n}; r = \overline{1,t+1}$ have finite variance. For f_0, we take the function of X_{t+1} we want to forecast.

We use the notation:

$$\mu_p(\theta) = E\lfloor f_p(X_r) \mid \theta \rfloor \tag{1.1}$$

$\left(p = \overline{0,n}; r = \overline{1,t+1}\right)$

This expression does not depend on r.

For this model we define the following structure parameters:

$$m_p = E\lfloor \mu_p(\theta) \rfloor = E\{E\lfloor f_p(X_r) \mid \theta \rfloor\} = E\lfloor f_p(X_r) \rfloor \quad (1.2),$$

$$a_{pq} = E\{Cov\lfloor f_p(X_r), f_q(X_r) \mid \theta \rfloor\} \quad (1.3),$$

$$b_{pq} = Cov\lfloor \mu_p(\theta), \mu_q(\theta) \rfloor \quad (1.4),$$

$$c_{pq} = Cov\lfloor f_p(X_r), f_q(X_r) \rfloor \quad (1.5),$$

$$d_{pq} = Cov\lfloor f_p(X_r), \mu_q(\theta) \rfloor \quad (1.6),$$

for $p, q = \overline{0,n}$. These expressions do not depend on $r = \overline{1,t+1}$. The structure parameters are connected by the following relations:

$$c_{pq} = a_{pq} + b_{pq} \quad (1.7),$$

$$d_{pq} = b_{pq} \quad (1.8),$$

for $p, q = \overline{0,n}$. This follows from the covariance relations obtained in the probability theory where they are very well-known.

Just as in the case of considering linear combinations of the observable variables themselves, we can also obtain non-homogeneous credibility estimates, taking as estimators the class of linear combinations of given functions of the observable variables, as shown in the following theorem:

Theorem 1.1 (Optimal non-homogeneous linearized estimators)

The linear combination of 1 and the random variables $f_p(X_r), p=\overline{1,n}; r=\overline{1,t}$ closest to $\mu_0(\theta)=E[f_0(X_{t+1})\,|\,\theta]$ and to $f_0(X_{t+1})$ in the least squares sense equals:

$$M=\sum_{p=1}^{n} z_p \sum_{r=1}^{t} \frac{1}{t} f_p(X_r)+m_0-\sum_{p=1}^{n} z_p m_p \qquad (1.9),$$

where $z_1, z_2, ..., z_n$ is a solution to the linear system of equations:

$$\sum_{p=1}^{n}\left[c_{pq}+(t-1)d_{pq}\right]z_p=td_{0q}\ (q=\overline{1,n}) \qquad (1.10)$$

or of the equivalent linear system of equations:

$$\sum_{p=1}^{n}\left(a_{pq}+tb_{pq}\right)z_p=tb_{0q}\ (q=\overline{1,n}) \qquad (1.11)$$

Applications of *Theorem 1.1*:

1) Let us consider the case of one given function f_1 in order to approximate $\mu_0(\theta)$. We formulate the following theorem. So for the special when n = 1, *Theorem 1.1* reads:

Theorem 1.2 (Optimal non-homogeneous linearized estimator, n = 1)

The linear combination of 1 and the random variables $f_1(X_r)$ ($r=\overline{1,t}$) closest to $\mu_0(\theta)$ in the least squares sense equals:

$$M=z\sum_{r=1}^{t}\frac{1}{t}f_1(X_r)+m_0-zm_1 \qquad (1.12)$$

where $m_1=E[f_1(X_r)],\ r=\overline{1,t},\ z=td_{01}/[c_{11}+(t-1)d_{11}]$ (1.13)

with $d_{01}=Cov[f_0(X_r), f_1(X_{r'})],\ d_{11}=Cov[f_1(X_r), f_1(X_{r'})],\ (r, r'=\overline{1,t}, r\neq r');$

$c_{11}=Cov[f_1(X_r), f_1(X_r)]$. (1.14)

2) The estimator M for $\mu_0(\theta)$ of *Theorem 1.1* can be displayed as:

$$M=f(X_1)+...+f(X_t) \qquad (1.15)$$

, where $f(x)=\frac{1}{t}\sum_{p=1}^{n}z_p f_p(x)+\frac{1}{t}m_0-\frac{1}{t}\sum_{p=1}^{n}z_p m_p$.

Let us forget now about this structure of f and look for any function f such that (1.15) is closest to $\mu_0(\theta)$. If are considered only functions f such that $f(X_1)$ has finite variance, then the optimal approximating function f results from the following theorem:

Theorem 1.3 (*Optimal approximating function or unique optimal function f*)

$f(X_1)+...+f(X_t)$ is closest to $\mu_0(\theta)$ and to $f_0(X_{t+1})$ in the least squares sense, if and only if f is a solution of the equation:

$$f(X_1)+(t-1)E[f(X_2)|X_1]-E[f_0(X_2)|X_1]\equiv 0 \tag{1.16}$$

Proof: we have to solve the following minimization problem:

$$\underset{g}{Min}\, E\{[f_0(X_{t+1})-g(X_1)-...-g(X_t)]^2\} \tag{1.17}$$

Suppose that f denotes the solution to this problem, then we consider: $g(X)=f(X)+\alpha h(X)$, with $h(\cdot)$ arbitrary, like in variational calculus. Let:

$$\varphi(\alpha)=E\{[f_0(X_{t+1})-f(X_1)-...-f(X_t)-\alpha h(X_1)-...-\alpha h(X_t)]^2\} \tag{1.18}$$

Clearly for f to be optimal, $\varphi'(0)=0$, so for every choice of h:

$$E\{[f_0(X_{t+1})-f(X_1)-...-f(X_t)][h(X_1)+...+h(X_t)]\}=0 \qquad (1.19),$$

, must hold. This can be rewritten as:

$$E[tf_0(X_2)h(X_1)-tf(X_1)h(X_1)-t(t-1)f(X_2)h(X_1)]=0 \qquad (1.20),$$

, or:

$$E[h(X_1)\{-f(X_1)-(t-1)E[f(X_2)|X_1]+E[f_0(X_2)|X_1]\}]=0 \tag{1.21}$$

Because this equation has to be satisfied for every choice of the function h one obtains, the expression in brackets in (1.21) must be identical to zero, which proves (1.16).

The following **example** is an **application** of *Theorem 1.3*:

If $X_1,...,X_{t+1}$ can only take the values $0,1,...,n$ and $p_{qr}=P[X_1=q,X_2=r]$ for: q, $r=\overline{0,n}$, then $f(X_1)+...+f(X_t)$ is closest to $\mu_0(\theta)$ and to $f_0(X_{t+1})$ in the least squares sense, if and only if for $q=\overline{0,n}$, $f(q)$ is a solution of the linear system:

$$f(q)\sum_{r=0}^{n}p_{qr}+(t-1)\sum_{r=0}^{n}f(r)p_{qr}=\sum_{r=0}^{n}f_0(r)p_{qr} \tag{1.22}$$

Indeed:

$$f(X_1):\begin{pmatrix} f(q) \\ P(X_1=q)\end{pmatrix}=\begin{pmatrix} f(q) \\ \sum_{r=0}^{n}p_{qr}\end{pmatrix},\ q=\overline{0,n};\ E[f(X_2)|X_1]=\sum_{r=0}^{n}f(r)P(X_2=r|X_1=$$

$$=q)=\sum_{r=0}^{n}f(r)\frac{p_{qr}}{\sum_{r=0}^{n}p_{qr}};\ E[f_0(X_2)|X_1]=\sum_{r=0}^{n}f_0(r)P(X_2=r|X_1=q)=\sum_{r=0}^{n}f_0(r)\frac{p_{qr}}{\sum_{r=0}^{n}p_{qr}}$$

Inserting these expressions for: $f(X_1)$, $E[f(X_2)|X_1]$ and $E[f_0(X_2)|X_1]$ into (1.16) leads to (1.22).

3) The special hierarchical semi-linear credibility theory is another **application** of the special semi-linear credibility theory.

Like in Jewell's hierarchical model we consider a portfolio of contracts, which can be broken up into P sectors each sector p consisting of k_p groups of contracts. Instead of estimating: $X_{p,j,t+1}$, $\mu(\theta_p,\theta_{pj}) = E[X_{p,j,t+1}|\theta_p,\theta_{pj}]$ (the pure net risk premium of the contract (p,j)), $\nu(\theta_p) = E[X_{p,j,t+1}|\theta_p]$ (the pure net risk premium of the sector p), we now estimate: $f_0(X_{p,j,t+1})$, $\mu_0(\theta_p,\theta_{pj}) = E[f_0(X_{p,j,t+1})|\theta_p,\theta_{pj}]$ (the pure net risk premium of the contract (p,j)), $\nu_0(\theta_p) = E[f_0(X_{p,j,t+1})|\theta_p]$ (the pure net risk premium of the sector p), where $p=\overline{1,P}$ and $j=\overline{1,k_p}$. In semi-linear credibility theory the following class of estimators is considered: $\alpha_0 + \sum_{p=1}^{n}\sum_{q=1}^{P}\sum_{i=1}^{k_q}\sum_{r=1}^{t}\alpha_{pqir} f_p(X_{qir})$, where $f_1(\cdot),\ldots,f_n(\cdot)$ are functions given in advance. Let us consider the case of one given function f_1 in order to approximate $f_0(X_{p,j,t+1})$ or $\nu_0(\theta_p)$ and $\mu_0(\theta_p,\theta_{pj})$. We formulate the following theorem:

Theorem 1.4 (*Hierarchical semi-linear credibility*)

Using the same notations as introduced for the hierarchical model of Jewell and denoting $X^0_{pjs} = f_0(X_{pjs})$ and $X^1_{pjs} = f_1(X_{pjs})$ one obtains the following least squares estimates for the pure net risk premiums:

$$\hat{\nu}_0(\theta_p) = (m_0 - z_p m_1) + z_p X^1_{pzw},\quad \hat{\mu}_0(\theta_p,\theta_{pj}) = (m_0 - z_{pj} m_1) + z_{pj} X^1_{pjw} \tag{3.1}$$

where: $X^1_{pjw} = \sum_{r=1}^{t}\frac{w_{pjr}}{w_{pj.}}X^1_{pjr}$, $X^1_{pzw} = \sum_{j=1}^{k_p}\frac{z_{pj}}{z_{p.}}X^1_{pjw}$, $z_{pj} = w_{pj.}d_{01}/[c_{11} + (w_{pj.}-1)d_{11}]$ (the credibility factor on contract level), with: $d_{01} = Cov(X^0_{pjr}, X^1_{pjr'})$, $d_{11} = Cov(X^1_{pjr}, X^1_{pjr'})$, $r \neq r'$, $c_{11} = Cov(X^1_{pjr}, X^1_{pjr}) = Var(X^1_{pjr})$, and: $z_p = z_{p.}D_{01}/[C_{11} + (z_{p.}-1)D_{11}]$ (the credibility factor at sector level), with: $D_{01} = Cov(X^0_{pjw}, X^1_{pj'w})$, $D_{11} = Cov(X^1_{pjw}, X^1_{pj'w})$, $j \neq j'$, $C_{11} = Cov(X^1_{pjw}, X^1_{pjw}) = Var(X^1_{pjw})$.

Remark 1.1 -the linear combination of 1 and the random variables X^1_{pjr} ($p=\overline{1,P}$, $j=\overline{1,k_p}$, $r=\overline{1,t}$) closest to $f_0(X_{p,j,t+1})$ and to $\nu_0(\theta_p)$ in the least squares sense equals $\hat{\nu}_0(\theta_p)$, and the linear combination of 1 and the random variables X^1_{pjr} ($p=\overline{1,P}$, $j=\overline{1,k_p}$, $r=\overline{1,t}$) closest to $\mu_0(\theta_p,\theta_{pj})$ in the least squares sense equals $\hat{\mu}_0(\theta_p,\theta_{pj})$.

Remark 1.2 -it should be noted that the solution (1.9) to the linearized credibility problem only yields a statistics computable from the observations, if the structure parameters are known. Generally, however, the structure function $U(\cdot)$ is not known. Then the "estimator" as it stands is not a statistic. Its interest is merely theoretical, but it will be the basis for further results on semi-linear credibility. In the following section we consider different contracts, each with the same structure parameters, so we can estimate these quantities using the statistics of the different contracts.

Section 3
Parameter estimation

The estimator obtained in the previous section contained structure parameters. In this section we assume the structure parameters are unknown, so the expressions for these (pseudo-) estimators are no longer statistics. But since the contracts are embedded in a collective of identical contracts, we now have more than one observation available on the risk parameter θ, so we can replace the unknown structure parameters by estimates. So now that we embedded the separate contract j in a collective of identical contracts, it is possible to give unbiased estimators of these quantities. It should be noted that the approximation to $f_0(X_{t+1})$ or to $\mu_0(\theta)$ based on a unique optimal approximating function f is always better than the one furnished in **Section 1** based on prescribed approximating functions $f_1, f_2, ..., f_n$. The usefulness of the latter approximation is that it is easy to apply, since it is sufficient to know estimates for the parameters a_{pq}, b_{pq} appearing in the credibility factors z_p. In this section we give some unbiased estimators for the parameters. For this purpose we consider k contracts, $j = \overline{1,k}$ and $k(\geq 2)$ independent and identically distributed random vectors $\left(\theta_j, \underline{X}_j^{'}\right) = \left(\theta_j, X_{j1}, ... X_{jt}\right)$, for $j = \overline{1,k}$. The contract indexed j is a random vector consisting of a random structure parameter θ_j and observations $X_{j1}, X_{j2}, ..., X_{jt}$, where $j = \overline{1,k}$. For every contract $j = \overline{1,k}$ and for θ_j fixed, the variables $X_{j1}, X_{j2}, ..., X_{jt}$ are conditionally independent and identically distributed. Here we will only derive estimators for the following parameters:

$$m_0 = E\left[f_0\left(X_{jr}\right)\right] \tag{2.1}$$

$$a_{01} = E\left\{Cov\left[f_0\left(X_{jr}\right), f_1\left(X_{jr}\right) \mid \theta_j\right]\right\} \tag{2.2}$$

$$b_{01} = Cov\left\{E\left[f_0\left(X_{jr}\right) \mid \theta_j\right], E\left[f_1\left(X_{jr}\right) \mid \theta_j\right]\right\} \tag{2.3}$$

One can prove the following theorem to hold.

Theorem 2.1 (Unbiased estimators for structure parameters)

Let:

$$\hat{m}_0 = \frac{1}{kt} X_{..}^0 = \frac{1}{kt} \sum_{j=1}^{k} \sum_{r=1}^{t} f_0(X_{jr}) \tag{2.4}$$

$$\hat{a}_{01} = \frac{1}{k(t-1)} \sum_{j=1}^{k} \sum_{r=1}^{t} \left(X_{jr}^0 - \frac{1}{t} X_{j.}^0 \right)\left(X_{jr}^1 - \frac{1}{t} X_{j.}^1 \right) \tag{2.5}$$

$$\hat{b}_{01} = \frac{1}{k-1} \sum_{j=1}^{k} \left(\frac{1}{t} X_{j.}^0 - \frac{1}{kt} X_{..}^0 \right)\left(\frac{1}{t} X_{j.}^1 - \frac{1}{kt} X_{..}^1 \right) - \frac{\hat{a}_{01}}{t} \tag{2.6}$$

then:

$$E\left(\hat{m}_0\right) = m_0,\ E\left(\hat{a}_{01}\right) = a_{01},\ E\left(\hat{b}_{01}\right) = b_{01} \tag{2.7}$$

We close this section giving as **application** of this estimation the unbiased estimators obtained for the structural parameters of the general semi-linear credibility model.

An **application** of *Theorem 2.1*:

The estimators

$$\hat{m}_p = \frac{1}{kt} X_{..}^p = \frac{1}{kt} \sum_{j=1}^{k} \sum_{r=1}^{t} f_p\left(X_{jr}\right) \tag{2.8}$$

$$\hat{a}_{pq} = \frac{1}{k(t-1)} \sum_{j=1}^{k} \sum_{r=1}^{t} \left(X_{jr}^p - \frac{1}{t} X_{j.}^p \right)\left(X_{jr}^q - \frac{1}{t} X_{j.}^q \right) \tag{2.9}$$

$$\hat{b}_{pq} = \frac{1}{k-1} \sum_{j=1}^{k} \left(\frac{1}{t} X_{j.}^p - \frac{1}{kt} X_{..}^p \right)\left(\frac{1}{t} X_{j.}^q - \frac{1}{kt} X_{..}^q \right) - \frac{\hat{a}_{pq}}{t} \tag{2.10}$$

are unbiased estimators of the corresponding structure parameters, i.e.:

$$E\left[\hat{m}_p\right] = m_p,\ E\left[\hat{a}_{pq}\right] = a_{pq},\ E\left[\hat{b}_{pq}\right] = b_{pq} \tag{2.11}$$

Indeed, we have:

$$E\left(\hat{m}_p\right) = \frac{1}{kt} \sum_{j=1}^{k} \sum_{r=1}^{t} E\left[f_p\left(X_{jr}\right)\right] = \frac{1}{kt} \sum_{j=1}^{k} \sum_{r=1}^{t} m_p = \frac{kt}{kt} m_p = m_p \tag{2.12}$$

(see (2.13)). So the verification of the first equality (2.11) is readily performed.

Remark 2.1. Note that the usual definitions of the structure parameters apply, with θ_j replacing θ and X_{jr} replacing X_r, so:

$$m_p = E\left[\mu_p\left(\theta_j\right)\right] = E\left\{E\left[f_p\left(X_{jr}\right) \mid \theta_j\right]\right\} = E\left[f_p\left(X_{jr}\right)\right] \tag{2.13}$$

$$a_{pq} = E\left\{Cov\left[f_p\left(X_{jr}\right), f_q\left(X_{jr}\right) \mid \theta_j\right]\right\} \tag{2.14}$$

$$b_{pq} = Cov\left[\mu_p\left(\theta_j\right), \mu_q\left(\theta_j\right)\right] = Cov\left\{E\left[f_p\left(X_{jr}\right) \mid \theta_j\right], E\left[f_q\left(X_{jr}\right) \mid \theta_j\right]\right\} \tag{2.15}$$

$$c_{pq} = Cov\left[f_p\left(X_{jr}\right), f_q\left(X_{jr}\right)\right] \tag{2.16}$$

Next:

$$E\left(\hat{a}_{pq}\right) = \frac{1}{k(t-1)} \sum_{j=1}^{k} \sum_{r=1}^{t} E\left[\left(X_{jr}^p - \frac{1}{t} X_{j.}^p \right)\left(X_{jr}^q - \frac{1}{t} X_{j.}^q \right)\right] =$$

$$=\frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}\left[Cov\left(X_{jr}^{p},X_{jr}^{q}\right)+E\left(X_{jr}^{p}\right)E\left(X_{jr}^{q}\right)-Cov\left(X_{jr}^{p},\frac{1}{t}X_{j.}^{q}\right)-\right.$$

$$-E\left(X_{jr}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right)-Cov\left(\frac{1}{t}X_{j.}^{p},X_{jr}^{q}\right)-E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(X_{jr}^{q}\right)+$$

$$\left.+Cov\left(\frac{1}{t}X_{j.}^{p},\frac{1}{t}X_{j.}^{q}\right)+E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right)\right] \quad (2.17)$$

But: $Cov\left(X_{jr}^{p},X_{jr'}^{q}\right)=Cov\left[f_p\left(X_{jr}\right),f_q\left(X_{jr'}\right)\right]=$

$$=\begin{cases} b_{pq} & (see(2.18)) \quad r\neq r' \\ a_{pq}+b_{pq} & (see(2.19)) \quad r=r' \end{cases}$$, because:

- for $r\neq r'$, we have:

$$Cov\left[f_p\left(X_{jr}\right),f_q\left(X_{jr'}\right)\right]=E\left[Cov\left(f_p\left(X_{jr}\right),f_q\left(X_{jr'}\right)\middle|\theta_j\right)\right]+Cov\left[E\left(f_p\left(X_{jr}\right)\middle|\theta_j\right),\right.$$

$$\left.,E\left(f_q\left(X_{jr'}\right)\middle|\theta_j\right)\right]=E\left\{E\left[f_p\left(X_{jr}\right)f_q\left(X_{jr'}\right)\middle|\theta_j\right]-E\left[f_p\left(X_{jr}\right)\middle|\theta_j\right]E\left[f_q\left(X_{jr'}\right)\middle|\theta_j\right]\right\}+$$

$$+Cov\left[\mu_p\left(\theta_j\right),\mu_q\left(\theta_j\right)\right]=E\left\{E\left[f_p\left(X_{jr}\right)\middle|\theta_j\right]E\left[f_q\left(X_{jr'}\right)\middle|\theta_j\right]-E\left[f_p\left(X_{jr}\right)\middle|\theta_j\right]E\left[f_q\left(X_{jr'}\right)\right.\right.$$

$$\left.\left.|\,\theta_j\right]\right\}+b_{pq}=b_{pq} \quad (2.18)$$

- for r = r’, we have:

$$Cov\left[f_p\left(X_{jr}\right),f_q\left(X_{jr}\right)\right]=E\left\{Cov\left[f_p\left(X_{jr}\right),f_q\left(X_{jr}\right)\middle|\theta_j\right]\right\}+Cov\left[E\left(f_p\left(X_{jr}\right)\middle|\theta_j\right),E\left(f_q\left(X_{jr}\right)\right.\right.$$

$$\left.\left.|\,\theta_j\right)\right]=a_{pq}+b_{pq} \quad (2.19)$$

So:

$$Cov\left(X_{jr}^{p},X_{jr'}^{q}\right)=\delta_{rr'}a_{pq}+b_{pq} \quad (2.20)$$

Next:

$$E\left(X_{jr}^{p}\right)=E\left[f_p\left(X_{jr}\right)\right]=m_p \quad (2.21)$$

$$E\left(X_{jr}^{q}\right)=E\left[f_q\left(X_{jr}\right)\right]=m_q \quad (2.22)$$

Also, we have:

$$Cov\left(X_{jr}^{p},\frac{1}{t}X_{j.}^{q}\right)=\frac{1}{t}\sum_{r'=1}^{t}Cov\left(X_{jr}^{p},X_{jr'}^{q}\right)=\frac{1}{t}\sum_{r'=1}^{t}\left(\delta_{rr'}a_{pq}+b_{pq}\right)=\frac{1}{t}a_{pq}+b_{pq}$$

So:

$$Cov\left(X_{jr}^{p},\frac{1}{t}X_{j.}^{q}\right)=\frac{1}{t}a_{pq}+b_{pq} \quad (2.23)$$

Next:

$$\mathrm{E}\left(\frac{1}{t}X_{j.}^{q}\right)=\frac{1}{t}\sum_{r=1}^{t}E\left(X_{jr}^{q}\right)=\frac{1}{t}\sum_{r=1}^{t}m_q=\frac{1}{t}tm_q=m_q \quad (2.24)$$

$$Cov\left(\frac{1}{t}X_{j.}^{p},X_{jr}^{q}\right)=\frac{1}{t}\sum_{r'=1}^{t}Cov\left(X_{jr'}^{p},X_{jr}^{q}\right)=\frac{1}{t}\sum_{r'=1}^{t}\left(\delta_{rr'}a_{pq}+b_{pq}\right)=\ldots=\frac{1}{t}a_{pq}+b_{pq}$$

(see the calculations from (2.23)).

So:

$$Cov\left(\frac{1}{t}X_{j.}^{p}, X_{jr}^{1}\right) = \frac{1}{t}a_{pq} + b_{pq} \tag{2.25}$$

Next:

$$\mathrm{E}\left(\frac{1}{t}X_{j.}^{p}\right) = \frac{1}{t}\sum_{r=1}^{t}E\left(X_{jr}^{p}\right) = \frac{1}{t}\sum_{r=1}^{t}m_p = \frac{1}{t}tm_p = m_p \tag{2.26}$$

and:

$$Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) = \frac{1}{t^2}\sum_{r=1}^{t}\sum_{r'=1}^{t}Cov\left(X_{jr}^{p}, X_{jr'}^{q}\right) = \frac{1}{t^2}\sum_{r=1}^{t}\sum_{r'=1}^{t}\left(\delta_{rr'}a_{pq} + b_{pq}\right) = \frac{1}{t}a_{pq} + b_{pq}$$

So:

$$Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) = \frac{1}{t}a_{pq} + b_{pq} \tag{2.27}$$

Inserting (2.20), (2.21), (2.22), (2.23), (2.24), (2.25), (2.26) and (2.27) in (2.17) one obtains:

$$E\left(\hat{a}_{pq}\right) = \frac{1}{k(t-1)}\sum_{j=1}^{k}\sum_{r=1}^{t}\Big[\left(a_{pq} + b_{pq}\right) + m_p m_q - \left(\frac{1}{t}a_{pq} + b_{pq}\right) - m_p m_q -$$

$$-\left(\frac{1}{t}a_{pq} + b_{pq}\right) - m_p m_q + \left(\frac{1}{t}a_{pq} + b_{pq}\right) + m_p m_q\Big] = a_{pq}$$

as was to be proven (see (2.11)). Finally, we have:

$$E\left(\hat{b}_{pq}\right) = \frac{1}{k-1}\sum_{j=1}^{k}\left[Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) + E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right) - Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{kt}X_{..}^{q}\right)\right.$$

$$- E\left(\frac{1}{t}X_{j.}^{p}\right)E\left(\frac{1}{kt}X_{..}^{q}\right) - Cov\left(\frac{1}{kt}X_{..}^{p}, \frac{1}{t}X_{j.}^{q}\right) - E\left(\frac{1}{kt}X_{..}^{p}\right)E\left(\frac{1}{t}X_{j.}^{q}\right) +$$

$$\left. + Cov\left(\frac{1}{kt}X_{..}^{p}, \frac{1}{kt}X_{..}^{q}\right) + E\left(\frac{1}{kt}X_{..}^{p}\right)E\left(\frac{1}{kt}X_{..}^{q}\right)\right] - \frac{a_{pq}}{t} \tag{2.28}$$

But:

$$Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{t}X_{j.}^{q}\right) = \frac{1}{t}a_{pq} + b_{pq} \tag{2.29}$$

(see (2.27))

Next:

$$Cov\left(\frac{1}{t}X_{j.}^{p}, \frac{1}{kt}X_{..}^{q}\right) = \frac{1}{kt^2}\sum_{r=1}^{t}\sum_{j'=1}^{k}\sum_{r'=1}^{t}Cov(X_{jr}^{p}, X_{j'r'}^{q}) = \frac{1}{kt^2}\sum_{r=1}^{t}\sum_{j'=1}^{k}\sum_{r'=1}^{t}\delta_{jj'}\left(\delta_{rr'}a_{pq} + b_{pq}\right)$$

$= \frac{1}{kt}a_{pq} + \frac{1}{k}b_{pq}$, because:

$$Cov\left(X_{jr}^{p}, X_{j'r'}^{q}\right) = \begin{cases} Cov\left(X_{j'r}^{p}, X_{j'r'}^{q}\right) & j = j' \\ Cov\left(X_{jr}^{p}, X_{j'r'}^{q}\right) & j \neq j' \end{cases} =$$

$$= \begin{cases} \delta_{rr'}a_{pq} + b_{pq} & (see(2.20)) \quad j = j' \\ 0 & (see(2.31)) \quad j \neq j' \end{cases} = \delta_{jj'}\left(\delta_{rr'}a_{pq} + b_{pq}\right) \tag{2.30}$$

where:

$Cov\left(X_{jr}^{p}, X_{j'r'}^{q}\right) = E\left[Cov\left(X_{jr}^{p}, X_{j'r'}^{q} \mid \theta_j\right)\right] + Cov\left[E\left(X_{jr}^{p} \mid \theta_j\right), E\left(X_{j'r'}^{q} \mid \theta_j\right)\right] =$

$= E\left[E\left(X_{jr}^{p} X_{j'r'}^{q} \mid \theta_j\right) - E\left(X_{jr}^{p} \mid \theta_j\right) E\left(X_{j'r'}^{q} \mid \theta_j\right)\right] + Cov\left[E\left(X_{jr}^{p} \mid \theta_j\right), E\left(X_{j'r'}^{q}\right)\right] =$

$= E[E\left(X_{jr}^{p} \mid \theta_j\right) E\left(X_{j'r'}^{q}\right) - E\left(X_{jr}^{p} \mid \theta_j\right) E\left(X_{j'r'}^{q}\right)] + 0 = E(0) = 0$, if $j \neq j'$.

So:

$$Cov\left(X_{jr}^{p}, X_{j'r'}^{q}\right) = 0 \quad (j \neq j') \tag{2.31}$$

and:

$$Cov\left(\frac{1}{t} X_{j.}^{p}, \frac{1}{kt} X_{..}^{q}\right) = \frac{1}{kt} a_{pq} + \frac{1}{k} b_{pq} \tag{2.32}$$

Next:

$$Cov\left(\frac{1}{kt} X_{..}^{p}, \frac{1}{t} X_{j.}^{q}\right) = \frac{1}{kt^2} \sum_{j'=1}^{k} \sum_{r=1}^{t} \sum_{r'=1}^{t} Cov\left(X_{j'r}^{p}, X_{jr'}^{q}\right) = \frac{1}{kt^2} \sum_{j'=1}^{k} \sum_{r=1}^{t} \sum_{r'=1}^{t} \delta_{jj'} \left(\delta_{rr'} a_{pq} + b_{pq}\right) =$$

$$= \frac{1}{kt} a_{pq} + \frac{1}{k} b_{pq}$$

(see (2.30)).

So:

$$Cov\left(\frac{1}{kt} X_{..}^{p}, \frac{1}{t} X_{j.}^{q}\right) = \frac{1}{kt} a_{pq} + \frac{1}{k} b_{pq} \tag{2.33}$$

Next:

$$Cov\left(\frac{1}{kt} X_{..}^{p}, \frac{1}{kt} X_{..}^{q}\right) = \frac{1}{k^2 t^2} \sum_{j=1}^{k} \sum_{r=1}^{t} \sum_{j'=1}^{k} \sum_{r'=1}^{t} Cov\left(X_{jr}^{p}, X_{j'r'}^{q}\right) = \frac{1}{k^2 t^2} \sum_{j=1}^{k} \sum_{r=1}^{t} \sum_{j'=1}^{k} \sum_{r'=1}^{t} \delta_{jj'} \left(\delta_{rr'}\right.$$

$$\left. a_{pq} + b_{pq}\right) = \frac{1}{kt} a_{pq} + \frac{1}{k} b_{pq}$$

So:

$$Cov\left(\frac{1}{kt} X_{..}^{p}, \frac{1}{kt} X_{..}^{q}\right) = \frac{1}{kt} a_{pq} + \frac{1}{k} b_{pq} \tag{2.34}$$

Also, we have:

$$E\left(\frac{1}{t} X_{j.}^{p}\right) = m_p \tag{2.35}$$

(see (2.26)).

$$E\left(\frac{1}{t} X_{j.}^{q}\right) = m_q \tag{2.36}$$

(see (2.24)).

$$E\left(\frac{1}{kt} X_{..}^{p}\right) = \frac{1}{kt} E\left(\sum_{j=1}^{k} \sum_{r=1}^{t} X_{jr}^{p}\right) = \frac{1}{kt} \sum_{j,r} E\left(X_{jr}^{p}\right) = \frac{1}{kt} ktm_p = m_p \tag{2.37}$$

(see (2.21)).

$$E\left(\frac{1}{kt}X_{..}^{q}\right)=\frac{1}{kt}E\left(\sum_{j=1}^{k}\sum_{r=1}^{t}X_{jr}^{q}\right)=\frac{1}{kt}\sum_{j,r}E\left(X_{jr}^{q}\right)=\frac{1}{kt}ktm_{q}=m_{q} \tag{2.38}$$

(see (2.22)). Inserting the values of the covariances and of the expectations (see (2.29), (2.32), (2.33), (2.34), (2.35), (2.36), (2.37) and (2.38) in (2.28), provides us with the desired results.

Indeed:

$$E\left(\hat{b}_{pq}\right)=\frac{1}{k-1}\sum_{j=1}^{k}\left[\left(\frac{1}{t}a_{pq}+b_{pq}\right)+m_pm_q-\left(\frac{1}{kt}a_{pq}+\frac{1}{k}b_{pq}\right)-m_pm_q-\left(\frac{1}{kt}a_{pq}+\right.\right.$$

$$\left.\left.+\frac{1}{k}b_{pq}\right)-m_pm_q+\left(\frac{1}{kt}a_{pq}+\frac{1}{k}b_{pq}\right)+m_pm_q\right]-\frac{a_{pq}}{t}=b_{pq}$$

as was to be proven (see (2.11)).

4. Conclusions

This paper completes the solution of the semi-linear credibility model in case of a non homogeneous linear estimator for $f_0(X_{j,t+1})$, or what amounts to the same, for $\mu_0(\theta_j)$. In view of the assumption about the independence of contracts, it might come as a surprise that the premium for 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 semi-linear credibility estimate for contract j. In this article, the semi-linear credibility model is refined by the introduction of the isolated contract j in a collective of contracts, all providing independent information on the structure distribution. But since the contracts are embedded in a collective of identical contracts, we now have more than one observation available on the risk parameter θ, so we can estimate these structural parameters in the semi-linear credibility model using the statistics of the different contracts. The above two theorems show that it is possible to give unbiased estimators of these quantities (the portfolio characteristics), if we embed the separate contract j in a collective of identical contracts. The article contains a description of the semi-linear credibility model, behind a heterogeneous 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 expectations and of conditional covariances become useful. The original model involving only one contract contains the basics of all further semi-linear credibility models. In the refined semi-linear credibility model a portfolio of contracts is studied, to be able to use the semi-linear credibility results.

Therefore, the main purpose of this paper 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, conditional covariances and variational calculus, needs not bother the user more than it does when he applies statistical tools like discriminatory analysis, scoring models, SAS and GLIM. These techniques can be applied by anybody on his own field of endeavor, be it economics, medicine, or insurance.

REFERENCES

[1]. Atanasiu V., *Contributions to the credibility theory*, Doctoral Dissertation, University of Bucharest-Faculty of Mathematics, 2000.
[2]. Atanasiu V., *Un model de credibilitate,* Revista Studii şi Cercetări de Calcul Economic şi Cibernetică Economică, XXXII, nr.3, 1998.
[3]. Goovaerts M.J., Kaas R., Van Heerwaarden, Bauwelinckx T., *Insurance Series,* volume 3, *Effective Actuarial Methods,* University of Amsterdam, The Netherlands, 1991.
[4]. Pentikäinen T., Daykin C.D.,Pesonen M., *Practical Risk Theory for Actuaries,* Université Pierré et Marie Curie, 1990.
[5]. Sundt B., *An Introduction to Non-Life Insurance Mathematics,* Veröffentlichungen des Instituts für Versicherungswissenschaft der Universität Mannheim Band 28, VVW Karlsruhe, 1990.