D I S C U S S I O N
P A P E R

0422

LINEAR CREDIBILITY MODELS BASED ON TIME SERIES FOR CLAIM COUNTS

O. PURCARU, M. GUILLÉN and M. DENUIT

LINEAR CREDIBILITY MODELS BASED ON TIME SERIES FOR CLAIM COUNTS

Oana Purcaru
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
purcaru@stat.ucl.ac.be

Montserrat Guillén
Dept. Econometrics
University of Barcelona
Diagonal, 690
S-08034 Barcelona, Spain
guillen@eco.ub.es

Michel Denuit
Institut de Statistique
Université Catholique de Louvain
Voie du Roman Pays, 20
B-1348 Louvain-la-Neuve, Belgium
denuit@stat.ucl.ac.be

September 10, 2004

Abstract

In this paper, several models for longitudinal analysis of claim frequencies are presented. These models are fitted to a large panel data set relating to a major Spanish insurance company. Credibility theory is then used to update the individual claim frequencies. The results are compared and the consequencies of increasing the strength of dependence are clearly visible.

Key words and phrases: GEE, frequency credibility, motor insurance, positive dependence, mixed Poisson model.

1 Introduction and Motivation

In this paper, we resort to three simple models for time series of claim counts, namely: Poisson-LogNormal with AR(1) structure in the random effects (on the log-scale), Poisson-LogNormal with exchangeable random effects and Poisson-LogNormal with static (commonotonic) random effects. These models are fitted to a large Spanish third part liability automobile data set, and their relative merits are discussed.

In the AR(1) case, the autocorrelation function decreases exponentially with time. Therefore, past claims are forgotten more rapidly than in the two other models, for which the autocorrelation function is constant. The latter models do not incorporate the age of the claim in risk prediction. We examine carefully the pattern of a posteriori corrections generated by the three models and emphasize the consequences of their application on the distribution of premiums among policyholders of the portofolio. It is important to mention that all the models possess the desirable financial stability property.

Let us now briefly describe the content of this paper. Section 2 presents the data set used to illustrate the work. Section 3 discusses the modelling of series of claim counts, including the estimation of the parameters. It also displays the estimations for the data of Section 2. Section 4 recalls the linear credibility procedure and carefully compares the values obtained for different types of policyholders, for claims of different ages and for the three models presented in Section 3. The final Section 5 concludes.

Before going further, let us make precise the notation that will be used throughout this paper. Random variables (vectors) will be denoted by capital (bold) letters. Normal distribution will be denoted as $\mathcal{N}(\mu, \sigma^2)$ and LogNormal distribution as $\mathcal{LN}(\mu, \sigma^2)$

2 Description of the data set

The data used in this paper are a ten percent sample of the automobile portofolio of a major insurance company operating in Spain. We only considered private use cars in this sample. The panel data contain information since 1991 until 1998.
Our sample contains 80.994 policyholders that stay in the company for seven complete yearly periods.

We have 18 exogeneous variables that are kept in the panel plus the yearly number of accidents. For every policy we have the initial information at the beginning of the period and the total number of claims at fault that took place whithin this yearly period. Some exogeneous information does not change with time, such as the insured's date of birth, but other information may change, for example, the type of guarantees that are covered by the contract.

Table 2.1 contains the observed yearly frequency of claims at fault from the first to the seventh period, together with the maximum number of claims per policyholder. The average claim frequency is 6.9%.

Period	Frequency	Maximum
1	0,079	4
2	0,070	4
3	0,063	5
4	0,064	4
5	0,066	4
6	0,069	4
7	0,075	4
Total	0.069	5

Table 2.1: Frequency of claims

The exogeneous variables are the ones described in Table 2.2.

Variable	Description
v2 to v7	vj equals 1 if record relates to the jth period, otherwise it is 0
v9	equals 1 for women and 0 for men
v10	equals 1 when driving in urban area, 0 otherwise
v11	equals 1 when zone is medium risk (Madrid and Catalonia)
v12	equals 1 when zone is high risk (Northern Spain)
v13	equals 1 if the driving license is between 4 and 14 years old
v14	equals 1 if the driving license is 15 or more years old
v15	equals 1 if the client has been in the company between 3 and 5 years
v16	equals 1 if the client has been in the company for more than 5 years
v17	equals 1 if the insured is 30 years old or younger
v18	equals 1 if coverage includes comprehensive except fire
v19	equals 1 if coverage includes comprehensive (material damage and fire)
v20	equals 1 if power is larger than or equal to 5500cc

Table 2.2: Exogeneous variables

3 Modelling through random effects

3.1 Description of the model

For each policyholder $i = 1, \ldots, n$ and coverage period $t = 1, \ldots, T_i$, we have at our disposal explanatory variables summarized in a vector $\boldsymbol{x}_{it}$. The information contained in the $\boldsymbol{x}_{it}$'s partly explains the behavior of the N_{it}'s. Since the actuary does not have access to important risk factor, like aggressiveness behind the wheel, there remains some heterogeneity in the portofolio. The effect of the unknown characteristics relating to policyholder i during year t

is represented in the model by a random variable Θ_{it}. The sequences of annual numbers of claims $\boldsymbol{N}_i = (N_{i1}, N_{i2}, \ldots)$ are assumed to be independent and the components $N_{i1}, \ldots, N_{iT_i}$ are assumed to be independent given the sequence $\boldsymbol{\Theta}_i = (\Theta_{i1}, \Theta_{i2}, \ldots, \Theta_{iT_i})$ of random effects. The latent unobservable process $\boldsymbol{\Theta}_i$ characterizes the correlation structure of the N_{it}'s.

Now, the ith policy of the portfolio, $i = 1, 2, \ldots, n$, is represented by a double sequence $(\boldsymbol{\Theta}_i, \boldsymbol{N}_i)$ where $\boldsymbol{N}_i$ gathers the observable annual claim numbers and $\boldsymbol{\Theta}_i$ is a positive random vector with unit mean representing the unexplained heterogeneity. Specifically, the model is based on the following assumptions:

B1 given $\boldsymbol{\Theta}_i = \boldsymbol{\theta}_i$, the random variables N_{it}, $t = 1, 2, \ldots, T_i$, are independent and conform to the Poisson distribution with mean $\lambda_{it}\theta_{it}$, i.e.

$$\Pr[N_{it} = k | \Theta_{it} = \theta_{it}] = \exp(-\lambda_{it}\theta_{it}) \frac{(\lambda_{it}\theta_{it})^k}{k!}, \;\; k \in \mathbb{N};$$

where $\lambda_{it} = d_{it} \exp(\boldsymbol{\beta}^t \boldsymbol{x}_{it})$ and d_{it} is the risk exposure.

B2 at the portfolio level, the sequences $(\boldsymbol{\Theta}_i, \boldsymbol{N}_i)$, $i = 1, 2, \ldots, n$, are assumed to be independent. Moreover, the Θ_{it}'s are non-negative random variables with unit mean ($\mathbb{E}[\Theta_{it}] = 1$ for all i, t):

- the distribution of $\boldsymbol{\Theta}_i = (\Theta_{i1}, \Theta_{i2}, \ldots, \Theta_{iT_i})$ depends only on T_i;
- defining $T_{\max} = \max_i T_i$, $\boldsymbol{\Theta}_i =_d (\Theta_1, \ldots, \Theta_{T_i})$, the distribution of $(\Theta_1, \ldots, \Theta_{T_{\max}})$ is supposed to be stationary.

Henceforth, we denote by $G(\cdot)$ (resp. $g(\cdot)$) the common cumulative distribution function (resp. probability density function) of the Θ_{it}'s, for $t = 1, \ldots, T_i$. We suppose also that the squared random effects are integrable.

3.2 Multivariate Poisson-LogNormal distribution

3.2.1 Description of the model

The multivariate LogNormal mixture of independent Poisson distributions is a parametric class of multivariate count distributions supporting a rich correlation structure. In our context, the multivariate Poisson-LogNormal law is a natural candidate to describe the joint distribution of the $\boldsymbol{N}_i$'s. This multivariate counting distribution has been considered by Aitchinson & Ho (1989); see also Joe (1997). Specifically, we have

$$\Pr[\boldsymbol{N}_i = \boldsymbol{n}_i] = \int_{\theta_{i1} \in \mathbb{R}^+} \cdots \int_{\theta_{iT_i} \in \mathbb{R}^+} \left\{ \prod_{t=1}^{T_i} \exp(-\theta_{it}\lambda_{it}) \frac{(\theta_{it}\lambda_{it})^{n_{it}}}{n_{it}!} \right\} f_{\boldsymbol{\Theta}_i}(\boldsymbol{\theta}_i | \boldsymbol{\Sigma}) d\theta_{i1} \ldots d\theta_{iT_i} \quad (3.1)$$

where the density $f_{\boldsymbol{\Theta}_i}$ corresponds to the multivariate $\mathcal{LN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ distribution, i.e.

$$f_{\boldsymbol{\Theta}_i}(\boldsymbol{\theta}_i | \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{T_i/2} \theta_{i1} \ldots \theta_{iT_i} |\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2} (\ln \boldsymbol{\theta}_i - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\ln \boldsymbol{\theta}_i - \boldsymbol{\mu}) \right\}$$

with $\boldsymbol{\mu}$ the mean vector and $\boldsymbol{\Sigma}$ the covariance matrix of the Gaussian random vector $\ln \boldsymbol{\Theta}_i$. In our context of serial dependence, $\boldsymbol{\Sigma} = \{\sigma_{s,t}\}_{1\leq s,t\leq T_i}$ has the form

$$\sigma_{tt} = \sigma^2 \text{ and } \sigma_{s,t} = \sigma(|t-s|) \text{ for } s \neq t,$$

and

$$\boldsymbol{\mu} = -\frac{1}{2}\sigma^2\mathbf{1}.$$

We conventionally put $\sigma(0) = \sigma^2$ for convenience. All the $\ln \Theta_{it}$'s conform thus to the same $\mathcal{N}(-\frac{1}{2}\sigma^2, \sigma^2)$ distribution.

3.2.2 First- and second-order moments

Although the multiple integral giving $\Pr[\boldsymbol{N}_i = \boldsymbol{n}_i]$ cannot be simplified, the moments of $\boldsymbol{N}_i$ are easily obtained in function of those of $\boldsymbol{\Theta}_i$. Let us first compute the moments of $\boldsymbol{\Theta}_i$:

$$\begin{aligned}
\mathbb{E}[\Theta_{it}] &= 1 \\
\mathbb{V}\text{ar}[\Theta_{it}] &= \mathbb{E}[\Theta_{it}^2] - \{\mathbb{E}[\Theta_{it}]\}^2 \\
&= \exp(\sigma^2) - 1 \\
\mathbb{C}\text{ov}[\Theta_{it}, \Theta_{is}] &= \mathbb{E}[\Theta_{it}\Theta_{is}] - \mathbb{E}[\Theta_{it}]\mathbb{E}[\Theta_{is}] \\
&= \exp\big(\sigma(|t-s|)\big) - 1.
\end{aligned}$$

Let us now compute the means and variances of the N_{it}'s:

$$\begin{aligned}
\mathbb{E}[N_{it}] &= \lambda_{it}\mathbb{E}[\Theta_{it}] = \lambda_{it} && (3.2) \\
\mathbb{V}\text{ar}[N_{it}] &= \mathbb{E}\Big[\mathbb{V}\text{ar}[N_{it}|\Theta_{it}]\Big] + \mathbb{V}\text{ar}\Big[\mathbb{E}[N_{it}|\Theta_{it}]\Big] \\
&= \lambda_{it} + \lambda_{it}^2\mathbb{V}\text{ar}[\Theta_{it}] \\
&= \lambda_{it}\Big\{1 + \lambda_{it}\{\exp(\sigma^2) - 1\}\Big\}. && (3.3)
\end{aligned}$$

For $s \neq t$, the covariance between N_{it} and N_{is} is given by

$$\begin{aligned}
\mathbb{C}\text{ov}[N_{it}, N_{is}] &= \mathbb{E}\Big[\mathbb{C}\text{ov}[N_{it}, N_{is}|\boldsymbol{\Theta}_i]\Big] + \mathbb{C}\text{ov}\Big[\mathbb{E}[N_{it}|\boldsymbol{\Theta}_i], \mathbb{E}[N_{is}|\boldsymbol{\Theta}_i]\Big] \\
&= \lambda_{it}\lambda_{is}\mathbb{C}\text{ov}[\Theta_{it}, \Theta_{is}] \\
&= \lambda_{it}\lambda_{is}\Big\{\exp\big(\sigma(|t-s|)\big) - 1\Big\}. && (3.4)
\end{aligned}$$

From (3.2)-(3.3), we see that the marginals of $\boldsymbol{N}_i$ have overdispersion relative to the Poisson distribution, as expected from any Poisson mixture. From (3.4) we have a correspondance between positive and negative values of correlations for the random effects Θ_{it} and the counts N_{it}.

3.2.3 Autocorrelation function of the latent process

Let us now introduce the autocorrelation function ρ_Θ, defined as

$$\rho_\Theta(h) = \mathbb{C}\text{orr}[\Theta_{it}, \Theta_{i,t+h}] = \frac{\exp\big(\sigma(h)\big) - 1}{\exp(\sigma^2) - 1}, \quad h = 1, 2, \ldots. \qquad (3.5)$$

This function features the temporal dependence between the random effects. If the predictive power of the claims decreases with their age, $\rho_\Theta(h)$ should decrease with h.

The decreasingness of ρ_Θ is in line with the exogeneous interpretation of residual heterogeneity (generated by unobservable risk characteristics). The correlation between rating factors relating to different periods should decrease with the lag, for observable variables as well as for hidden features. Similarly, endogeneous components of latent processes (e.g. effort variables modeling moral hazard) yield the same conclusion: a driver will tend to be more prudent after an accident but this effect certainly decreases with the age of the claim.

Computing

$$\begin{aligned}\mathbb{C}\mathrm{orr}[N_{it}, N_{is}] &= \frac{\mathbb{C}\mathrm{ov}[N_{it}, N_{is}]}{\sqrt{\mathbb{V}\mathrm{ar}[N_{it}]\mathbb{V}\mathrm{ar}[N_{is}]}} \\ &= \frac{\sqrt{\lambda_{it}\lambda_{is}}\Big\{\exp\big(\sigma(|t-s|)\big)-1\Big\}}{\sqrt{1+\lambda_{it}\{\exp(\sigma^2)-1\}}\sqrt{1+\lambda_{is}\{\exp(\sigma^2)-1\}}} \text{ for } s \neq t.\end{aligned}$$

shows that

$$\mathbb{C}\mathrm{orr}[N_{it}, N_{it+h}] = \frac{\rho_\Theta(h)}{\sqrt{1+\frac{1}{\lambda_{it}\{\exp(\sigma^2)-1\}}}\sqrt{1+\frac{1}{\lambda_{it+h}\{\exp(\sigma^2)-1\}}}}, \quad h \geq 1.$$

Note that the inequality

$$\mathbb{C}\mathrm{orr}[N_{it}, N_{it+h}] \leq \mathbb{C}\mathrm{orr}[\Theta_{it}, \Theta_{it+h}]$$

is true, so that the correlation is always decreased by mixing.

3.2.4 Estimating the parameters

In order to estimate the model parameters, a direct likelihood maximisation is extremely difficult since the likelihood $\prod_{i=1}^n \Pr[\boldsymbol{N}_i = \boldsymbol{n}_i]$ with $\Pr[\boldsymbol{N}_i = \boldsymbol{n}_i]$ of the form (3.1) has no explicit form. It is worth to mention that methods have been developed for Poisson-LogNormal count processes (as the Monte-Carlo approach of Chan & Ledolter (1995)), but seem to be out of reach for the very large data sets the actuaries have to treat. Therefore, we prefer a GEE approach, which is in line with Poisson-LogNormal model, except that it is only based upon first- and second-order moments. Since all estimators are consistent, we do not expect much difference in practice. The advantage of the GEE approach over the classical Poisson regression (assuming serial independence, whose estimators are also consistent) is that the standard errors of the regression parameters are corrected.

For convenience, let us consider a panel data set (T observation periods for each policyholder with $T = 7$ for the Spanish data used in the numerical illustrations). Let us define the $T \times T$ diagonal matrix $\boldsymbol{A}_i$ as

$$\boldsymbol{A}_i = \begin{pmatrix} \lambda_{i1} & 0 & \cdots & 0 \\ 0 & \lambda_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_{iT} \end{pmatrix}.$$

We also introduce $\boldsymbol{S}_i = \boldsymbol{N}_i - \boldsymbol{\lambda}_i$ and $\boldsymbol{X}_i$ as the part of the design matrix relating to policyholder i (i.e. $\boldsymbol{X}_i$ is of dimension $T \times p$, where p is the number of observed covariates, and has lines $\boldsymbol{x}_{i1}^t, \ldots, \boldsymbol{x}_{iT}^t$). We denote as $\boldsymbol{\alpha}$ the vector of parameters relating to the dependence structure; we chose

$$\alpha_h = \mathbb{C}\mathrm{ov}[\Theta_t, \Theta_{t+h}]$$

but other parameterizations are of course possible.

It is well-known that the maximum likelihood equation for $\boldsymbol{\beta}$ in the Poisson model with serial independence (i.e. if $N_{i1}, \ldots, N_{iT_i}$ are mutually independent) is

$$\sum_{i=1}^{n} \boldsymbol{X}_i^t \boldsymbol{S}_i = \boldsymbol{0}. \tag{3.6}$$

The solution $\widehat{\boldsymbol{\beta}}^{\perp}$ of (3.6) remains consistent if the elements of $\boldsymbol{N}_i$ are correlated (provided the λ_{it}'s are correctly specified).

Let us now present estimating equations taking the correlation into account to increase efficiency. Since the estimator of $\boldsymbol{\beta}$ resulting from (3.6) remains consistent, we do not expect much difference for point estimation of the regression coefficients. Accounting for serial correlation inflates the standard errors of the estimations of the β_j's, which is important, e.g. in the context of variable selection as well as to correctly assess the accuracy of the pure premium estimation.

Let $\boldsymbol{D}_i = \boldsymbol{A}_i \boldsymbol{X}_i$. We can view (3.6) as the sum of $\boldsymbol{D}_i^t \boldsymbol{A}_i^{-1} \boldsymbol{S}_i$, where A_i is the covariance matrix of $\boldsymbol{N}_i$ in the model with serial independence. The idea is now to substitute the covariance matrix $\boldsymbol{V}_i$ of $\boldsymbol{N}_i$ in the Poisson mixture model to $\boldsymbol{A}_i$, and to estimate $\boldsymbol{\beta}$ by the solution of

$$\sum_{i=1}^{n} \boldsymbol{D}_i^t \boldsymbol{V}_i^{-1} \boldsymbol{S}_i = \boldsymbol{0}. \tag{3.7}$$

Zeger & Liang (1986) showed that this estimator possesses the desirable property of consistency.

3.3 Application to Spanish panel data

3.3.1 Model 1 - Static Poisson-LogNormal distribution

In the static LogNormal model we have the random effects $\boldsymbol{\Theta}_i = (\Theta_i, \ldots, \Theta_i)$ with each Θ_i conforming to the LogNormal distribution (for more details see Purcaru & Denuit (2001)). Note that in this case $\rho_\Theta(h) = 1$ for all h.
In this model, the random effect Θ_{it}'s representing the unobservable characteristics does not change over time: it is given from the beginning. Consequently, the dependence between the N_{it}'s is very strong, since they all share the same $\boldsymbol{\Theta}_i$.
The estimators of the parameters are displayed in Table 3.1. Step1 is obtained by solving (3.6). An interative procedure is then used to get the solution of (3.7). Convergence occured rapidly (6 step were needed to reach stability). Table 3.2 displays confidence intervals for the regression parameters. When the estimations are performed under the independence assumption, all the β's are significantly different from 0 (since no confidence interval overlaps

0). On the contrary, the confidence interval for β_1 contains 0 when serial dependence is taken into account. This is a typical consequence of serial correlation: the explanatory power of exogeneous variables is overestimated under the falsely assumed independence hypothesis.

estimators	step1	step2	step3	step4	step5	step6
$\hat{\beta}_0$	-2.262888	-2.328807	-2.327847	-2.327869	-2.327868	-2.327868
$\hat{\beta}_1$	0.0362971	0.0365472	0.0361616	0.0361683	0.0361679	0.0361679
$\hat{\beta}_2$	-0.047071	-0.044836	-0.044519	-0.044518	-0.044518	-0.044518
$\hat{\beta}_3$	-0.051469	-0.055623	-0.055187	-0.05518	-0.05518	-0.05518
$\hat{\beta}_4$	0.1832375	0.1826864	0.1827176	0.1827119	0.1827119	0.1827119
$\hat{\beta}_5$	-0.366918	-0.344407	-0.345133	-0.345117	-0.345118	-0.345118
$\hat{\beta}_6$	-0.423503	-0.393731	-0.394259	-0.394245	-0.394245	-0.394245
$\hat{\beta}_7$	-0.136436	-0.120047	-0.120234	-0.120227	-0.120227	-0.120227
$\hat{\beta}_8$	-0.216371	-0.170516	-0.170598	-0.17059	-0.17059	-0.17059
$\hat{\beta}_9$	0.1006956	0.1086738	0.1081477	0.108147	0.1081469	0.1081469
$\hat{\beta}_{10}$	0.1802757	0.2033404	0.2037509	0.2037408	0.2037411	0.2037411
$\hat{\beta}_{11}$	0.083799	0.0884854	0.0884599	0.0884547	0.0884548	0.0884548
$\hat{\beta}_{12}$	0.1048213	0.1037882	0.1033855	0.1033932	0.1033929	0.1033929
$\hat{\sigma}^2_\Theta$	1.3761605	1.3845667	1.3840476	1.3840431	1.3840429	1.3840429

Table 3.1: Static - LogNormal model - estimations of the parameters

estimators	IC (independence hypothesis)		IC (dependence hypothesis)	
$\hat{\beta}_0$	-2.332428	-2.193347	-2.409273	-2.246463
$\hat{\beta}_1$	0.0070609	0.0655334	-0.004998	0.0773341
$\hat{\beta}_2$	-0.069482	-0.02466	-0.076547	-0.012489
$\hat{\beta}_3$	-0.078063	-0.024875	-0.094152	-0.016209
$\hat{\beta}_4$	0.1571729	0.209302	0.1426775	0.2227463
$\hat{\beta}_5$	-0.421799	-0.312037	-0.407571	-0.282665
$\hat{\beta}_6$	-0.483288	-0.363717	-0.462165	-0.326325
$\hat{\beta}_7$	-0.170956	-0.101916	-0.155339	-0.085115
$\hat{\beta}_8$	-0.24842	-0.184322	-0.206682	-0.134499
$\hat{\beta}_9$	0.0659831	0.135408	0.068178	0.1481159
$\hat{\beta}_{10}$	0.1506472	0.2099041	0.1649785	0.2425037
$\hat{\beta}_{11}$	0.059629	0.107969	0.0563114	0.1205981
$\hat{\beta}_{12}$	0.0782991	0.1313434	0.068647	0.1381389

Table 3.2: Static - LogNormal model - 95% Confidence Intervals for $\hat{\boldsymbol{\beta}}$'s

3.3.2 Model 2 - AR(1) Poisson-LogNormal

A simple and powerful time series model for count data is obtained by specifying a Gaussian autoregressive process of order one (AR(1), in short) for the Θ_{it}'s on the log-scale, that is:

$$\ln \Theta_{it} = \varrho \ln \Theta_{it-1} + \epsilon_{it},$$

where $|\varrho| < 1$ and the ϵ_{it}'s are independent, $\mathcal{N}\big(-\frac{\sigma^2}{2}(1-\varrho), \sigma^2(1-\varrho^2)\big)$ errors (the variance ensures that the $\ln \Theta_{it}$'s remain $\mathcal{N}(-\frac{\sigma^2}{2}, \sigma^2)$ distributed.) For a reference in actuarial science, see Pinquet, Guillén and Bolancé (2001).

In the AR(1) Poisson-LogNormal model, $\sigma(h) = \sigma^2 \varrho^h$ so that from (3.5) and since $\sigma_\Theta^2 = \exp(\sigma^2) - 1$ it follows that the autocorrelation function is given by:

$$\rho_\Theta(h) = \frac{\exp(\sigma^2 \varrho^h) - 1}{\exp(\sigma^2) - 1} = \frac{(\sigma_\Theta^2 + 1)^{\varrho^h} - 1}{\sigma_\Theta^2}. \tag{3.8}$$

The estimations of the model parameters, are given in Table 3.3. It can be seen that the convergence is rather fast: the algorithm converged (the maximum change in the parameter estimates is less than 1E-6) after 5 iterations and that all of the obtained values are significant. Step1 represents the solution of (3.6). The 95% confidence intervals are presented in Table 3.4.

estimators	step1	step2	step3	step4	step5	step6
$\hat{\beta}_0$	-2.262888	-2.301964	-2.301219	-2.301236	-2.301235	-2.301235
$\hat{\beta}_1$	0.0362971	0.0386112	0.0383588	0.0383656	0.0383653	0.0383653
$\hat{\beta}_2$	-0.047071	-0.045409	-0.045277	-0.045277	-0.045277	-0.045277
$\hat{\beta}_3$	-0.051469	-0.053754	-0.053512	-0.053512	-0.053512	-0.053512
$\hat{\beta}_4$	0.1832375	0.1842641	0.1843198	0.1843171	0.1843171	0.1843171
$\hat{\beta}_5$	-0.366918	-0.343935	-0.344518	-0.344503	-0.344504	-0.344504
$\hat{\beta}_6$	-0.423503	-0.399743	-0.400348	-0.400332	-0.400333	-0.400333
$\hat{\beta}_7$	-0.136436	-0.133157	-0.133181	-0.13318	-0.13318	-0.13318
$\hat{\beta}_8$	-0.216371	-0.193163	-0.193268	-0.193268	-0.193268	-0.193268
$\hat{\beta}_9$	0.1006956	0.1048603	0.1046811	0.1046788	0.1046788	0.1046788
$\hat{\beta}_{10}$	0.1802757	0.1950995	0.1950677	0.195066	0.1950661	0.1950661
$\hat{\beta}_{11}$	0.083799	0.085373	0.0853426	0.0853404	0.0853405	0.0853405
$\hat{\beta}_{12}$	0.1048213	0.1047803	0.1046262	0.1046287	0.1046286	0.1046286
$\hat{\sigma}_\Theta^2$	1.3761605	1.3697761	1.3696191	1.3696244	1.3696243	1.3696243
$\hat{\varrho}$	0.8355049	0.8352565	0.8352557	0.8352556	0.8352556	0.8352556

Table 3.3: AR(1) - LogNormal model - estimations of the parameters

estimators	IC (serial independence hypothesis)		IC (serial dependence hypothesis)	
$\hat{\beta}_0$	-2.332428	-2.193347	-2.383616	-2.218854
$\hat{\beta}_1$	0.0070609	0.0655334	0.0002118	0.0765187
$\hat{\beta}_2$	-0.069482	-0.02466	-0.074788	-0.015766
$\hat{\beta}_3$	-0.078063	-0.024875	-0.088724	-0.0183
$\hat{\beta}_4$	0.1571729	0.209302	0.1484957	0.2201385
$\hat{\beta}_5$	-0.421799	-0.312037	-0.40921	-0.279798
$\hat{\beta}_6$	-0.483288	-0.363717	-0.47029	-0.330375
$\hat{\beta}_7$	-0.170956	-0.101916	-0.170113	-0.096247
$\hat{\beta}_8$	-0.24842	-0.184322	-0.230581	-0.155955
$\hat{\beta}_9$	0.0659831	0.135408	0.0639855	0.1453722
$\hat{\beta}_{10}$	0.1506472	0.2099041	0.1578262	0.2323061
$\hat{\beta}_{11}$	0.059629	0.107969	0.0547408	0.1159402
$\hat{\beta}_{12}$	0.0782991	0.1313434	0.0715963	0.1376608

Table 3.4: AR(1) - LogNormal model - 95% Confidence Intervals for $\hat{\boldsymbol{\beta}}$'s

As expected, taking serial dependence into account increases the standard deviation of the estimates and this makes the confidence intervals bigger compared to independence. Nevertheless, all the $\boldsymbol{\beta}$ values remain significant (at 5%).

3.3.3 Model 3 - Product Poisson-LogNormal

Let

$$\Theta_{it} = U_i W_{it},$$

where the W_{it}'s are independent and identically distributed, and independent from U_i, with

$$W_{it} =_d \text{LogNormal}(-\frac{\sigma_W^2}{2}, \sigma_W^2) \qquad \text{and} \qquad U_i =_d \text{LogNormal}(-\frac{\sigma_U^2}{2}, \sigma_U^2).$$

Then $\Theta_{it} =_d \text{LogNormal}(-\frac{\sigma_W^2+\sigma_U^2}{2}, \sigma_W^2 + \sigma_U^2)$.
In this model the covariance of the random effects depends only of the variance of the time-independent component, i.e.

$$\begin{aligned}
\mathbb{C}\text{ov}[\Theta_{it}, \Theta_{it+h}] &= \mathbb{E}\Big[\mathbb{C}\text{ov}[\Theta_{it}, \Theta_{it+h}|U_i]\Big] + \mathbb{C}\text{ov}\Big[\mathbb{E}[\Theta_{it}|U_i], \mathbb{E}[\Theta_{it+h}|U_i]\Big] \\
&= \mathbb{C}\text{ov}\Big[U_i\mathbb{E}[W_{it}], U_i\mathbb{E}[W_{it+h}]\Big] \\
&= \mathbb{V}\text{ar}[U_i] = \sigma_U^2,
\end{aligned}$$

Thus the autocorrelation function between the random effects is constant:

$$\rho_\Theta(h) \equiv \rho = \frac{\sigma_U^2}{\exp(\sigma^2) - 1} = \frac{\sigma_U^2}{\sigma_\Theta^2}, \quad \text{for} \quad h \geq 1. \tag{3.9}$$

The estimations of this model and their 95% confidence intervals are shown in Table 3.5 and Table 3.6, respectively. Again, convergence occurs very rapidly. Step1 is obtained by solving (3.6) for β. Since the variance gets inflated once serial dependence has been taken into account, confidence intervals become bigger, as it can be seen from Table 3.6.

estimators	step1	step2	step3	step4	step5	step6
$\hat{\beta}_0$	-2.262888	-2.317991	-2.317451	-2.317454	-2.317453	-2.317453
$\hat{\beta}_1$	0.0362971	0.0387336	0.0383974	0.038403	0.0384028	0.0384028
$\hat{\beta}_2$	-0.047071	-0.045623	-0.045387	-0.045386	-0.045386	-0.045386
$\hat{\beta}_3$	-0.051469	-0.054624	-0.054322	-0.054317	-0.054317	-0.054317
$\hat{\beta}_4$	0.1832375	0.1829461	0.1829939	0.1829897	0.1829898	0.1829898
$\hat{\beta}_5$	-0.366918	-0.348265	-0.348707	-0.348701	-0.348702	-0.348702
$\hat{\beta}_6$	-0.423503	-0.399244	-0.399566	-0.399562	-0.399562	-0.399562
$\hat{\beta}_7$	-0.136436	-0.123077	-0.123189	-0.123188	-0.123188	-0.123188
$\hat{\beta}_8$	-0.216371	-0.178943	-0.178895	-0.178897	-0.178897	-0.178897
$\hat{\beta}_9$	0.1006956	0.1075345	0.1072077	0.1072033	0.1072033	0.1072033
$\hat{\beta}_{10}$	0.1802757	0.198145	0.1985319	0.1985226	0.1985227	0.1985226
$\hat{\beta}_{11}$	0.083799	0.0875983	0.0876039	0.0875988	0.0875989	0.0875989
$\hat{\beta}_{12}$	0.1048213	0.1054846	0.1051544	0.1051596	0.1051595	0.1051595
$\hat{\sigma}^2_\Theta$	1.3761605	1.3828968	1.3825758	1.382571	1.382571	1.382571
$\hat{\rho}$	0.756513	0.7598022	0.7596325	0.7596295	0.7596295	0.7596295

Table 3.5: Product - LogNormal model - estimations of the parameters

estimators	IC (independence hypothesis)		IC (dependence hypothesis)	
$\hat{\beta}_0$	-2.332428	-2.193347	-2.398503	-2.236403
$\hat{\beta}_1$	0.0070609	0.0655334	-0.000794	0.0775995
$\hat{\beta}_2$	-0.069482	-0.02466	-0.075789	-0.014983
$\hat{\beta}_3$	-0.078063	-0.024875	-0.090944	-0.017689
$\hat{\beta}_4$	0.1571729	0.209302	0.1455656	0.2204139
$\hat{\beta}_5$	-0.421799	-0.312037	-0.411699	-0.285704
$\hat{\beta}_6$	-0.483288	-0.363717	-0.467918	-0.331206
$\hat{\beta}_7$	-0.170956	-0.101916	-0.158974	-0.087403
$\hat{\beta}_8$	-0.24842	-0.184322	-0.215096	-0.142699
$\hat{\beta}_9$	0.0659831	0.135408	0.0672727	0.1471339
$\hat{\beta}_{10}$	0.1506472	0.2099041	0.1609052	0.2361401
$\hat{\beta}_{11}$	0.059629	0.107969	0.0565795	0.1186183
$\hat{\beta}_{12}$	0.0782991	0.1313434	0.0716146	0.1387045

Table 3.6: Product - LogNormal model - 95% Confidence Intervals for $\hat{\boldsymbol{\beta}}$'s

3.3.4 Comparison of frequency premiums for the three models

Let us compare the result between the two models emphasizing the time-dependence and the one for the time-independence.
We have listed the estimated regression coefficients in Table 3.8, but these values are difficult to compare. For this reason we have identified three policyholders of the portofolio: the worst driver, the best driver and the average driver. Those are described in Table 3.7. The estimated expected claim frequency according to the three models are given in Table 3.8.

Policyholder	v9	v10	v11	v12	v13	v14	v15	v16	v17	v18	v19	v20
GOOD	0	1	1	0	0	1	0	1	0	0	0	0
AVERAGE	0	1	0	1	1	0	1	0	0	0	1	1
BAD	1	0	0	1	0	0	0	0	1	1	0	1

Table 3.7: Risk classes in the portofolio

	Models		
estimators	AR(1) - LogNormal	Product - LogNormal	Static - LogNormal
$\hat{\beta}_0$	-2.301235	-2.317453	-2.327868
$\hat{\beta}_1$	0.0383653	0.0384028	0.0361679
$\hat{\beta}_2$	-0.045277	-0.045386	-0.044518
$\hat{\beta}_3$	-0.053512	-0.054317	-0.05518
$\hat{\beta}_4$	0.1843171	0.1829898	0.1827119
$\hat{\beta}_5$	-0.344504	-0.348702	-0.345118
$\hat{\beta}_6$	-0.400333	-0.399562	-0.394245
$\hat{\beta}_7$	-0.13318	-0.123188	-0.120227
$\hat{\beta}_8$	-0.193268	-0.178897	-0.17059
$\hat{\beta}_9$	0.1046788	0.1072033	0.1081469
$\hat{\beta}_{10}$	0.1950661	0.1985226	0.2037411
$\hat{\beta}_{11}$	0.0853405	0.0875989	0.088454
$\hat{\beta}_{12}$	0.1046286	0.1051595	0.1033929
		Premiums	
Good	0.0501055	0.0500059	0.0501668
Average	0.0863014	0.0855221	0.0851644
Bad	0.1874620	0.1854120	0.1838367

Table 3.8: Comparison of frequency premiums for the three models

Clearly, the different models give very similar point estimates for the expected claim frequencies. Nevertheless the three models will yield very different credibility premiums.

4 Comparison of linear credibility updating formulas

4.1 Derivation of linear credibility formulas

In the spirit of Bühlmann, linear credibility predictors are obtained from a linear regression derived in the model with random effects. This is particularly appealing in our context since ZEGER's (1988) extension of the Poisson regression model allowing for serial correlation is distributionally known only up to the first- and second-order moments.
Let $\mathcal{H}_{it}$ be the claim history of the policyholder up to (and including) time $t-1$ (that is $N_{i1}, \ldots, N_{i,t-1}$ together with frequency premiums $\lambda_{i1}, \ldots, \lambda_{it-1}$). The predictor of $\mathbb{E}[\Theta_{i,T_i+1}|\mathcal{H}_{i,T_i}]$ is of the form

$$a_i + \sum_{t=1}^{T_i} b_{it} n_{it} \tag{4.1}$$

where n_{it} is the number of claims reported by policyholder i during year t. The coefficients a_i and b_{it}, $t = 1, \ldots, T_i$, are determined in order to minimize the expected square difference between the unknown Θ_{i,T_i+1} and the linear credibility predictor (4.1), that is, to minimize

$$\psi(a_i, b_{i1}, \ldots, b_{iT_i}) = \mathbb{E}\left[\left(\Theta_{i,T_i+1} - a_i - \sum_{t=1}^{T_i} b_{it} N_{it}\right)^2\right].$$

Equating $\frac{\partial}{\partial a_i}\psi$ to zero yields

$$a_i + \sum_{t=1}^{T_i} b_{it}\lambda_{it} = 1$$

whence it follows that we have to minimize

$$\widetilde{\psi}(b_{i1}, \ldots, b_{iT_i}) = \mathbb{E}\left[\left\{(\Theta_{i,T_i+1} - 1) - \sum_{t=1}^{T_i} b_{it}(N_{it} - \lambda_{it})\right\}^2\right].$$

The stationary equation $\frac{\partial}{\partial b_{is}}\widetilde{\psi}$ yields

$$\mathbb{C}\mathrm{ov}[\Theta_{i,T_i+1}, N_{is}] - \sum_{t=1}^{T_i} b_{it}\mathbb{C}\mathrm{ov}[N_{is}, N_{it}] = 0 \tag{4.2}$$

where

$$\mathbb{C}\mathrm{ov}[\Theta_{i,T_i+1}, N_{is}] = \mathbb{C}\mathrm{ov}\Big[\Theta_{i,T_i+1}, \mathbb{E}[N_{is}|\boldsymbol{\Theta}_i]\Big] = \lambda_{is}\Big\{\exp\big(\sigma(T_i + 1 - s)\big) - 1\Big\}.$$

Solving the linear system (4.2) provides the b_{it}'s.

Clearly, the predictors (4.1) depend not only on past claims but also on past premiums (that is, on past characteristics of policyholders). Consequently, they average to one whatever the rating factors. This fairness property is not always fulfilled by experience rating schemes enforced in real-life: if the boni and the mali do not depend on *a priori* ratemaking (that is, on the frequency of claims), the average *a posteriori* corrections coefficients increase with the frequency. In order to illustrate the extent to which a posteriori corrections depend on a priori characteristics, let us consider three typical policyholders:

1. a bad driver, i.e. a policyholder paying high *a priori* premiums
2. an average driver, i.e. a policyholder paying average *a priori* premiums
3. a good driver, i.e. a policyholder paying low *a priori* risk premium.

Those persons are precisely described in Table 4.1. They are in line with the risks described in Tables 3.7 and 3.8. We consider that they pay 100% of the premium for being covered the first year.

Policyholder	*a priori* premium
Good	5%
Medium	10%
Bad	20%

Table 4.1: Risk classes

In the remainder of this section we will carefully examine:

(i) how the age of the claims is incorporated in a posteriori correction (see subsection 4.2)

(ii) how the observable characteristics of policyholder affect a posteriori correction (see subsection 4.3)

(iii) how the choice of the model for the series of claims counts influences a posteriori corrections (see subsection 4.4)

4.2 *A posteriori* correction according to age of claims

4.2.1 Model 1 - Static-LogNormal

Table 4.2 describes the evolution of the expected annual claim frequency for a good driver having reported a single claim in year 1 (column 2) or in year 5 (column 3). We observe that both scenarios yield exactly the same a posteriori corrections. This model does not take into account the age of claims: the predictive ability of a claim does not decrease with time, which is quite restrictive. Intuitively speaking, the first scenario should yield a lower claim frequency for year 10 than scenario 2.

	Year of claim	
Years	$y = 1$	$y = 5$
1	0.05	0.05
2	**0.17359**	0.04350
3	0.15361	0.03849
4	0.13775	0.03451
5	0.12486	0.03129
6	*0.11418*	**0.11418**
7	0.10518	0.10518
8	0.09749	0.09749
9	0.09085	0.09085
10	0.08506	0.08506

Table 4.2: GOOD policyholder : premiums assuming the Static - LogNormal model

4.2.2 Model 2 - AR(1)-LogNormal

Table 4.3 describes the evolution of the expected claim frequency relating to a good policyholder if he has reported a single claim during the first year (column 2) or during the 5-th year (column 3). The penalisation is seen in year 2 or year 6, respectively (in bold). Considering the premium of year 7, we clearly see how the age of the claims affects the amount paid by the policyholder : the amount in the second case by far exceeds the corresponding value in the first situation. A recent claim has thus more impact on the revised expected claim frequency than an old one.

	Year of claim	
Years	$y = 1$	$y = 5$
1	0.05	0.05
2	**0.074047**	0.048734
3	0.069817	0.047488
4	0.065997	0.046177
5	0.062392	0.044739
6	0.058871	**0.086875**
7	0.055327	0.080344
8	0.051668	0.074028
9	0.047798	0.067790
10	0.043615	0.061494

Table 4.3: GOOD policyholder : premiums assuming the AR(1) - LogNormal model

4.2.3 Model 3 - Product-LogNormal

Table 4.4 is similar to Table 4.3. In this case, the premiums for year 6 are identical in both situations. As Model 1, this model does not incorporate the age of claims in the risk prediction and this is due to the fact that ρ_Θ is constant.

	Year of claim	
Years	$y = 1$	$y = 5$
1	0.05	0.05
2	**0.14372**	0.04507
3	0.13082	0.04102
4	0.12004	0.03764
5	0.11090	0.03478
6	*0.10305*	**0.10305**
7	0.09625	0.09625
8	0.09028	0.09028
9	0.08501	0.08501
10	0.08033	0.08033

Table 4.4: GOOD policyholder : premiums assuming the Product - LogNormal model

In passing, we see that the a posteriori corrections are more severe for Model 1 than for Model 3. This is due to the fact that the dependence existing between the N_{it}'s is stronger in Model 1 than in Model 3 so that the credibility system penalizes more heavily past claim for Model 3.

4.3 *A posteriori* correction according to a priori characteristics

Assume either that the three individuals described in Table 4.1 (good, average and bad driver) report a single claim during the first year. The evolution of their expected claim frequency is presented in Tables 4.7, 4.9 and 4.5, according to the model retained to describe serial dependence. Tables 4.8, 4.10 and 4.6 are the analogues if those three kind of drivers do not report any claim.

4.3.1 Model 1 - Static-LogNormal

By observing the Table 4.5, we conclude that the relative increase of premium due to the reporting of one claim during the year 1 is the biggest for the good policyholder. On the whole, the lower the claim frequency, the higher the relative a posteriori correction is.

In Table 4.6, when there is no claim reported, we observe as in the two previous cases that the relative discounts granted to bad policyholder are the highest. On the whole, the higher the claim frequency, the higher the relative a posteriori discounts.

Year	Policyholder Good		Average		Bad	
1	0.05	100%	0.10	100%	0.20	100%
2	**0.17359**	347.18%	**0.30721**	307.21%	**0.49944**	249.72%
3	0.15361	307.21%	0.24972	249.72%	0.36341	181.71%
4	0.13775	275.50%	0.21035	210.35%	0.28562	142.81%
5	0.12486	249.72%	0.18171	181.71%	0.23526	117.63%
6	0.11418	228.35%	0.15993	159.93%	0.20	100%
7	0.10518	210.35%	0.14281	142.81%	0.17393	86.97%
8	0.09749	194.98%	0.12900	129.01%	0.15387	76.94%
9	0.09085	181.71%	0.11763	117.63%	0.13796	68.98%
10	0.08506	170.12%	0.10810	108.10%	0.12504	62.52%

Table 4.5: Static LogNormal model - premiums (values and % of the *a priori* premium) if a claim is filed during the 1-st year

Year	Policyholder Good		Average		Bad	
1	0.05000	0%	0.10000	0%	0.20000	0%
2	0.04350	13%	0.07698	23.02%	0.12514	37.43%
3	0.03849	23.02%	0.06257	37.43%	0.09106	54.47%
4	0.03452	30.96%	0.05271	47.29%	0.07157	64.22%
5	0.03129	37.42%	0.04553	54.47%	0.05895	70.53%
6	0.02861	42.78%	0.04007	59.93%	0.05011	74.95%
7	0.02635	47.3%	0.03578	64.22%	0.04358	78.21%
8	0.02443	51.14%	0.03232	67.68%	0.03855	80.73%
9	0.02277	54.46%	0.02947	70.53%	0.03457	82.72%
10	0.02131	57.38%	0.02709	72.91%	0.03133	84.34%

Table 4.6: Static LogNormal model - premiums and the associated discounts (in %) if no claims are reported over 10 years

In this case, this result straightly follows from the explicit expression

$$\mathbb{E}[\Theta_{i,T_i+1}|\mathcal{H}_{i,T_i}] = \frac{1+\sigma_\Theta^2\sum_{t=1}^{T_i} n_{it}}{1+\sigma_\Theta^2\sum_{t=1}^{T_i}\lambda_{it}},$$

which clearly shows how for fixed n_{it}'s the λ_{it}'s affect the a posteriori correction.

4.3.2 Model 2 - AR(1)-LogNormal

As it can be seen from line 2 in Table 4.7, the relative increase of premium due to the reporting of one claim during the year 1 is the biggest for good policyholder. On the whole, the lower the claim frequency, the higher the relative a posteriori correction.

In table 4.8 we see that the relative discounts granted to bad policyholders, when they do not report any claims is the highest. On the whole, the higher the claim frequency, the higher relative a posteriori discounts. The conclusions of Model 1 carry to Model 2, even if in the latter case there is no explicit expression for $\mathbb{E}[\Theta_{i,T_i+1}|\mathcal{H}_{i,T_i}]$.

	Policyholder					
Year	Good		Average		Bad	
1	0.05	100%	0.10	100%	0.20	100%
2	**0.074047**	148.09%	**0.14040**	140.40%	**0.25854**	129.27%
3	0.069817	139.63%	0.12770	127.70%	0.22641	113.20%
4	0.065997	131.99%	0.11734	117.34%	0.20313	101.56%
5	0.062392	124.79%	0.10817	108.17%	0.18383	91.92%
6	0.058871	117.74%	0.09958	99.58%	0.16635	83.18%
7	0.055327	110.66%	0.09116	91.16%	0.14933	74.67%
8	0.051668	103.34%	0.08256	82.56%	0.13176	65.88%
9	0.047798	95.60%	0.07348	73.48%	0.11275	56.38%
10	0.043615	87.23%	0.06360	63.60%	0.09148	45.74%

Table 4.7: AR(1)-LogNormal model - premiums (values and % of the *a priori* premium) if a claim is filed during the 1-st year

	Policyholder					
Year	Good		Average		Bad	
1	0.05	0%	0.1	0%	0.2	0%
2	0.048734	2.54%	0.09551	4.49%	0.18537	7.32%
3	0.047488	5.024%	0.09171	8.29%	0.17543	12.29%
4	0.046177	7.65%	0.088	12%	0.16643	16.79%
5	0.044739	10.522%	0.08407	15.93%	0.1571	21.45%
6	0.043117	13.77%	0.07972	20.28%	0.14679	26.61%
7	0.041255	17.49%	0.07477	25.23%	0.135	32.5%
8	0.03909	21.82%	0.06902	30.98%	0.12125	39.38%
9	0.036551	26.9%	0.06229	37.71%	0.10504	47.48%
10	0.033558	32.88%	0.05435	45.65%	0.08583	57.09%

Table 4.8: AR(1)-LogNormal model - premiums and the associated discounts (in %) if no claims are reported over 10 years

4.3.3 Model 3 - Product-LogNormal

Tables 4.9 and 4.10 reveal the same patterns for the a posteriori corrections acording to the the a priori penalty of the risk, i.e. the relative increase of premium due to the reporting of one claim during the year 1 is the biggest for good policyholder, and the relative discounts granted to bad policyholders, when they do not report any claims is the highest.

On the whole,the lower (higher) claim frequency, the higher the relative a posteriori correction (discount) is.

	Policyholder					
Year	Good		Average		Bad	
1	0.05	100%	0.10	100%	0.20	100%
2	**0.14372**	287.44%	**0.25717**	257.17%	**0.42718**	213.59%
3	0.13082	261.63%	0.21893	218.94%	0.33270	166.35%
4	0.12004	240.08%	0.19060	190.60%	0.27245	136.22%
5	0.11090	221.80%	0.16876	168.76%	0.23067	115.33%
6	0.10305	206.11%	0.15141	151.41%	0.20	100%
7	0.09625	192.49%	0.13729	137.29%	0.17653	88.27%
8	0.09028	180.56%	0.12558	125.58%	0.15799	78.99%
9	0.08501	170.03%	0.11572	115.72%	0.14297	71.49%
10	0.08033	160.65%	0.10729	107.29%	0.13056	65.28%

Table 4.9: Product-LogNormal model - premiums (values and % of the *a priori* premium) if a claim is filed during the 1-st year

	Policyholder					
Year	Good		Average		Bad	
1	0.05	0%	0.1	0%	0.2	0%
2	0.045067	9.87%	0.08254	17.46%	0.1432	28.4%
3	0.04102	17.96%	0.07027	29.73%	0.11153	44.24%
4	0.03764	24.72%	0.06117	38.83%	0.09133	54.34%
5	0.034775	30.45%	0.05416	45.84%	0.07733	61.34%
6	0.032315	35.37%	0.04859	51.41%	0.06705	66.48%
7	0.03018	39.64%	0.04406	55.94%	0.05918	70.41%
8	0.02831	43.38%	0.04031	59.69%	0.05296	73.52%
9	0.026658	46.68%	0.03714	62.86%	0.04793	76.04%
10	0.025188	49.624%	0.03443	65.57%	0.04377	78.12%

Table 4.10: Product-LogNormal model - premiums and the associated discounts (in %) if no claims are reported over 10 years

4.4 *A posteriori* correction according to the model used for series of claim counts

4.4.1 BAD policyholder

In Table 4.11 there is one feature which appear to be interesting, namely the time needed for the "remission of sins": how long does it take for the penalty to disappear. If there is a single claim reported during year 1, the initial level 0.2 is obtained once again after a period of 4 years, in the Product-LogNormal and the Static-LogNormal models. In the AR(1)-LogNormal model 2 years are enough to forget about the claim. A consequence of such phenomenon is that the discounts, when there is no claim reported, are smaller in the AR(1)-LogNormal model, than in the two others. Only the penalties obtained from the AR(1)-LogNormal model could be enforced in practice.

Claim	Models AR(1) LN	 Product LN	 Static LN	Year
	0.20000	0.20000	0.20000	1
	0.18537	0.14320	0.12514	2
	0.17543	0.11153	0.09106	3
	0.16643	0.09133	0.07157	4
no claims	0.15710	0.07733	0.05895	5
(over 10 years)	0.14679	0.06705	0.05011	6
	0.13500	0.05918	0.04358	7
	0.12125	0.05296	0.03855	8
	0.10504	0.04793	0.03457	9
	0.08583	0.04377	0.03133	10
	0.20000	0.20000	0.20000	1
	0.25854	0.42718	0.49944	2
	0.22641	0.33270	0.36341	3
	0.20313	0.27245	0.28562	4
1 claim	0.18383	0.23067	0.23526	5
(1-st year)	0.16635	0.20000	0.20000	6
	0.14933	0.17653	0.17393	7
	0.13176	0.15799	0.15387	8
	0.11275	0.15799	0.13796	9
	0.09148	0.13056	0.12504	10

Table 4.11: BAD policyholder - *a posteriori* correction

4.4.2 AVERAGE policyholder

The phenomenon of "remission of sins" for the average policyholder is no longer so obvious, as it was for the bad policyholder. However, we observe in table 4.12 that if there is a single claim reported in the first year, the premiums are decreasing faster (3 years) toward the premium paid in year 1 with the AR(1)-LogNormal model, than the two other models.

One consequence of this feature is that the discounts, when no claims are reported, in the AR(1)-LogNormal case are the smallest.

Claim	Models AR(1) LN	Product LN	Static LN	Year
no claims (over 10 years)	0.10000	0.10000	0.10000	1
	0.09551	0.08254	0.07698	2
	0.09171	0.07027	0.06257	3
	0.08800	0.06117	0.05271	4
	0.08407	0.05416	0.04553	5
	0.07972	0.04859	0.04007	6
	0.07477	0.04406	0.03578	7
	0.06902	0.04031	0.03232	8
	0.06229	0.03714	0.02947	9
	0.05435	0.03443	0.02709	10
1 claim (1-st year)	0.10000	0.10000	0.10000	1
	0.14040	0.25717	0.30721	2
	0.12770	0.21893	0.24972	3
	0.11734	0.19060	0.21035	4
	0.10817	0.16876	0.18171	5
	0.09958	0.15141	0.15993	6
	0.09116	0.13729	0.14281	7
	0.08256	0.12558	0.12900	8
	0.07348	0.11572	0.11763	9
	0.06360	0.10729	0.10810	10

Table 4.12: AVERAGE policyholder - *a posteriori* correction

4.4.3 GOOD policyholder

Table 4.13 is similar to Table 4.12. We observe that for the good policyholder if there is a single claim reported in the first year, the premiums are decreasing faster (6 years) toward the premium paid in year 1 with the AR(1)-LogNormal model, than the two other models.

As in the two previous cases, one consequence of this feature is that the discounts, when no claims are reported, in the AR(1)-LogNormal case are the smallest.

Claim	Models AR(1) LN	 Product LN	 Static LN	Year
	0.050000	0.050000	0.050000	1
	0.048734	0.045067	0.043495	2
	0.047488	0.041020	0.038488	3
	0.046177	0.037640	0.034515	4
no claims	0.044739	0.034775	0.031285	5
(over 10 years)	0.043117	0.032315	0.028608	6
	0.041255	0.030180	0.026353	7
	0.039090	0.028310	0.024428	8
	0.036551	0.026658	0.022765	9
	0.033558	0.025188	0.021313	10
	0.050000	0.05000	0.05000	1
	0.074047	0.14372	0.17359	2
	0.069817	0.13082	0.15361	3
	0.065997	0.12004	0.13775	4
1 claim	0.062392	0.11090	0.12486	5
(1-st year)	0.058871	0.10305	0.11418	6
	0.055327	0.09625	0.10518	7
	0.051668	0.09028	0.09749	8
	0.047798	0.08501	0.09085	9
	0.043615	0.08033	0.08506	10

Table 4.13: GOOD policyholder - *a posteriori* correction

5 Conclusions

The results of this paper illustrate the different kinds of a posteriori risk evaluation produced by the three models for series of claim counts considered in this study: Static-LogNormal, Product-LogNormal and AR(1)-LogNormal.

In the future, we aim to apply copula models to generate Markov of order one structure for the latent process. This will allow us to have more freedom concerning the marginal distribution of the random effects (in particular, gamma distributed latent variables are desirable).

Acknowledgements

The support of the Belgian Government under contract "Projet d'Actions de Recherche Concertées" ARC 98/03-217 is gratefully acknowledged.

References

[1] Aitchinson, J., and Ho, C.H. (1989). The multivariate Poisson-log normal distribution. *Biometrika* **75**, 621-629.

[2] Chan, K.S., & Ledolter, J. (1995). Monte Carlo EM estimation for time series models involving counts. *Journal of the American Statistical Association* **90**, 242-252.

[3] Joe, H. (1997). *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

[4] Pinquet, J., Guillen, M., & Bolancé, C. (2001). Allowance for the age of claims in bonus-malus systems. *ASTIN Bulletin*, **31**, 2, 337-348.

[5] Purcaru, O., & Denuit, M. (2002). On the dependence induced by frequency credibility models. *Belgian Actuarial Bulletin*, **2**, 1, 73-79.

[6] Purcaru, O., & Denuit, M. (2003). Dependence in dynamic claim frequency credibility models. *ASTIN*, **33**, 1, 23-40.

[7] Zeger, S.L. (1988). A regression model for time series of counts. *Biometrika* **75**, 621-629.

[8] Zeger, S.L., & Liang K-Y.(1986). Longitudinal data analysis using generalized linear models.*Biometrika* **73**, 13-22.