Greatest accuracy credibility with dynamic heterogeneity: the Harvey-Fernandes model

Catalina Bolancé[1], **Michel Denuit**[2], **Montserrat Guillén**[3] **and Philippe Lambert**[4]

Abstract. This paper proposes a generalization of the merit rating system described in Dionne & Vanasse (1989,1992). The model takes into account explanatory variables as well as possible modifications in the policyholders unobservable risk characteristics. It is based on the analysis of time series for count observations proposed by Harvey & Fernandes (1989). Numerical results obtained with a Spanish panel databasis for motor insurance illustrate the approach described in this paper.

Keywords: Poisson-Gamma mixture, credibility theory, a posteriori ratemaking, time series models for counts.

1 Introduction and Motivation

In many European and Asian countries, as well as in North American states or provinces, insurers relate premium amounts to individual past experience in motor insurance. More specifically, they penalize insured drivers responsible for one or more accidents by premium surcharges (or *maluses*) and reward claim-free policyholders by awarding them discounts (or *bonuses*).

Insurers also base risk classification on variables easy to observe, such as age, gender, type of vehicle, and so on. However, the tariff classes they obtain are not homogeneous and this may generate a ratemaking structure that is unfair to the insured drivers. In order to match individual's premium to risk and to increase incentives for road safety, the past record is then taken into consideration under a merit rating scheme. These two forms of ratemaking can be justified by asymmetrical information between the insurance company and the policyholders.

In this paper, we aim to construct a no-claim discount system in an empirical Bayesian framework: Bayesian analysis is used to compute the posterior distribution of the claim frequency for policyholders who experienced similar histories. Such a system is fair: Any insured has to pay at each renewal an amount of premium proportional to the estimate of his claim frequency taking into account, through Bayes theorem, all the information gathered in the past.

The claims history of each policyholder consists in a short integer-valued sequence of yearly claim counts. The basic model used for experience rating is based on the Negative Binomial distribution. This probability law can be seen as a Poisson mixture distribution with Gamma mixing. Therefore, it accomodates for serial dependence of claim counts, by introducing Gamma-distributed unobserved individual heterogeneity. The serial dependence in claim counts sequences is generated by integrating the unobserved factor, and by updating its prediction when individual information increases. Alternative models with LogNormal of Inverse Gaussian unobserved heterogeneity have also been considered in the actuarial literature.

The vast majority of the papers appeared in the actuarial literature considered time-independent heterogeneous models. The model considered in this paper includes an unknown underlying random parameter that develops over time: Instead of assuming that the risk characteristics are given once and for all by a single risk parameter, we suppose that the unknown risk characteristics of each policy are described by dynamic random effects. In the terminology of Jewell (1975), these are evolutionary credibility models, which means that the underlying risk parameter is allowed to vary in successive periods (in other words, the structure function is allowed to be time dependent). Evolutionary credibility models for claim amounts have been considered, e.g., in Gerber & Jones (1975), Jewell (1975,1976) and Sundt (1981,1983,1988). In this paper, following Albrecht (1985), we consider evolutionary, or dynamic, credibility models for claim numbers. We are in line with recent contributions by Pinquet, Guillén & Bolancé (2001), Bolancé, Pinquet & Guillén (2003) and Purcaru, Guillén & Denuit (2004).

Other approaches have also been considered in the literature. Gouriéroux & Jasiak (2004) have applied the integer valued autoregressive (so-called INAR) model to account for serial dependence in count processes. See also Brannas & Hellstrom (2001), Freeland & McCabe (2004) and the references therein. Gouriéroux & Jasiak (2004) also extended INAR models to accomodate for unobserved heterogeneity, typically present in motor insurance. Other models applied in medical studies could also be contemplated by actuaries. In that respect, let us mention the zero-inflated Poisson mixed autoregressive model for analyzing time series of count events with excess zeros. See Yau, Lee & Carrivick (2004) for an illustration and for useful references.

Let us now detail the contents of this paper. Section 2 de-

[1] Department of Econometrics, University of Barcelona, E-08034 Barcelona, Spain
[2] Institut de Sciences Actuarielles, Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium
[3] Department of Econometrics, University of Barcelona, E-08034 Barcelona, Spain
[4] Institut des Sciences Humaines et Sociales, Université de Liège, B-4000 Liège

scribes the time series model for count data proposed by Harvey & Fernandes (1989). Section 3 discusses the resulting merit-rating system. It will be seen there that we get credibility estimators with geometric weights, already encountered in the actuarial literature. The estimation of the dependence parameter will be carefully examined. Section 4 is devoted to a numerical illustration. Considering the Spanish databasis analyzed by Pinquet, Guillén & Bolancé (2001) and Bolancé, Pinquet & Guillén (2003), we apply the merit rating system designed in Section 3 to a panel of 80 894 policyholders followed from 1991 to 1997. The effects of the implementation of this experience rating mechanism are carefully examined, in particular the high bonus-hunger effect induced by the dynamic mixed Poisson model.

To end with, let us briefly present the notations used in this paper. We consider a portfolio consisting of n policies. For the ith policyholder, $i = 1, 2, \ldots, n$, let T_i be the number of years elapsed since this policy has been issued, and let N_{it}, $t = 1, 2, \ldots, T_i$, denote the number of claims reported by this policyholder during the t-th period of insurance. Let $\mathcal{P}oi(\theta)$, $\theta > 0$, denote the Poisson distribution with mean θ, i.e. $\mathcal{P}oi(\theta)$ has discrete probability density function $\exp(-\theta)\theta^k/k!$, $k = 0, 1, 2, \ldots$, and let $\mathcal{G}am(a, \tau)$, $a, \tau > 0$, denote the Gamma distribution with mean a/τ and variance a/τ^2, i.e. $\mathcal{G}am(a, \tau)$ has density function $\tau^a \exp(-\tau x) x^{a-1}/\Gamma(a)$, $x \geq 0$.

2 Harvey-Fernandes model

Suppose that we observe independent series of claim counts $\{N_{i1}, N_{i2}, \ldots, N_{iT_i}\}$, $i = 1, 2, \ldots, n$, on the n policies of a portfolio. Let $\mathcal{F}_{it}$ be the claim history for policy i up to time t. Formally, $\mathcal{F}_{it}$ is the sigma algebra generated by the random variables $N_{i1}, N_{i2}, \ldots, N_{it}$. Our purpose is to define a model for count data allowing the underlying mean of the process to change over time. Following Harvey & Fernandes (1989), we introduce an hyperparameter α discounting past observations in making forecasts: The predictions can be constructed by a sort of exponentially moving average procedure. For extensions of this model, see Lambert (1996a,b).

A common problem for count data is that, even after allowing for important explanatory variables using the Poisson regression model, the fits obtained are rather poor. This indicates that, conditional upon the explanatory variables included in the final model, the variance of an observation is greater than its mean, implying that the Poisson assumption is incorrect. Most often, this is due to the fact that important explanatory variables may not have been measured and are consequently incorrectly excluded from the regression relationship.

A convenient way to take this phenomenon into account is to introduce a random effect in this model. Let $\boldsymbol{x}_{it}$ be the vector of all the relevant covariates for policyholder i during year t (including age, gender and power of the car, for instance). Here, we consider the following Poisson-Gamma model:

$$[N_{it}|\Theta_{it}, \boldsymbol{\beta}, \mathcal{F}_{i;t-1}] \sim \mathcal{P}oi\big(\Theta_{it} \exp(\boldsymbol{\beta}'\boldsymbol{x}_{it})\big) \qquad (1)$$

where $[\Theta_{it}|\mathcal{F}_{it}] \sim \mathcal{G}am(a_{it}, \tau_{it})$, the parameters a_{it} and τ_{it} being computed from the first t observations. We require that the mean of $[\Theta_{it}|\mathcal{F}_{i,t-1}]$ is the same as that of $[\Theta_{i,t-1}|\mathcal{F}_{i,t-1}]$, but the variance increases. This effect is induced by multiplying the Gamma parameters by a factor α less than 1. We therefore suppose that

$$[\Theta_{it}|\mathcal{F}_{i,t-1}] \sim \mathcal{G}am(a_{i,t|t-1}, \tau_{i,t|t-1}),$$

with parameters $a_{i,t|t-1}$ and $\tau_{i,t|t-1}$ such that

$$\begin{cases} a_{i,t|t-1} = \alpha a_{i,t-1} \\ \tau_{i,t|t-1} = \alpha \tau_{i,t-1} \end{cases}$$

and $0 < \alpha \leq 1$. Then,

$$\mathbb{E}[\Theta_{it}|\mathcal{F}_{i,t-1}] = \frac{a_{i,t|t-1}}{\tau_{i,t|t-1}} = \frac{a_{i,t-1}}{\tau_{i,t-1}} = \mathbb{E}[\Theta_{i,t-1}|\mathcal{F}_{i,t-1}]$$

whereas

$$\mathbb{V}\mathrm{ar}[\Theta_{it}|\mathcal{F}_{i,t-1}] = \frac{a_{i,t|t-1}}{\tau^2_{i,t|t-1}} = \alpha^{-1}\mathbb{V}\mathrm{ar}[\Theta_{i,t-1}|\mathcal{F}_{i,t-1}].$$

The last formula expresses the type of correlation existing between the N_{it}'s for fixed i: values of claim counts during close periods are closely related, but the strength of this dependence decreases with time. This is the key formula in the Harvey-Fernandes model, in that it links the variances of the random effect before and after the addition of a new observation. The appropriateness of this proportionality assumption of course heavily depends on the data set under study.

Application of Bayes theorem then yields a posterior distribution for $[\Theta_{it}|\mathcal{F}_{it}]$ given by

$$d\mathbb{P}[\Theta_{it} \leq \theta|\mathcal{F}_{it}] \propto \exp\Big(-\theta\exp(\boldsymbol{\beta}'\boldsymbol{x}_{it})\Big)\frac{\big(\theta\exp(\boldsymbol{\beta}'\boldsymbol{x}_{it})\big)^{N_{it}}}{N_{it}!} \tau_{i,t|t-1}^{a_{i,t|t-1}}\frac{\exp(-\tau_{i,t|t-1}\theta)\theta^{a_{i,t|t-1}}}{\Gamma(a_{i,t|t-1})}.$$

3 Resulting merit rating system

Once the observation N_{it} becomes available, the posterior distribution of Θ_{it} is then given by the Gamma distribution with parameters

$$\begin{cases} a_{it} = \alpha a_{it-1} + N_{it} \\ \tau_{it} = \alpha \tau_{it-1} + \exp\big(\boldsymbol{\beta}'\boldsymbol{x}_{it}\big) \end{cases} \qquad (2)$$

The initial prior distribution (that is, the distribution of Θ_{i1}) is taken to be $\mathcal{G}am\,(a_{i0}, \tau_{i0})$, with $a_{i0} = \tau_{i0}$. Indeed, we then have $\mathbb{E}[\Theta_{i1}] = 1$ so that the premium for the first year results from *a priori* rating and equals $\exp\big(\boldsymbol{\beta}'\boldsymbol{x}_{i1}\big)$. Note that the a_{i0}'s are taken to be identical for all the policyholders in the portfolio, so that $a_{i0} = a_0$ for all i. The parameter a_0 reflects the residual heterogeneity of the portfolio. Repeated substitutions in (2) lead to

$$\begin{cases} a_{it} = \sum_{l=0}^{t-1} \alpha^l N_{i,t-l} + \alpha^t a_0 \\ \tau_{it} = \sum_{l=0}^{t-1} \alpha^l \exp\big(\boldsymbol{\beta}'\boldsymbol{x}_{it-l}\big) + \alpha^t a_0 \end{cases}$$

The expected annual claim frequency for year $t+1$ is

$$\begin{aligned}
&\mathbb{E}\left[N_{it+1}|\mathcal{F}_{it}\right]\\
&= \exp\left(\boldsymbol{\beta}'\boldsymbol{x}_{i,t+1}\right)\mathbb{E}\left[\Theta_{it+1}|\mathcal{F}_{it}\right]\\
&= \exp\left(\boldsymbol{\beta}'\boldsymbol{x}_{i,t+1}\right)\frac{a_{it}}{\tau_{it}}\\
&= \exp\left(\boldsymbol{\beta}'\boldsymbol{x}_{i,t+1}\right)\frac{a_0+\sum_{l=1}^{t}\alpha^{-l}N_{il}}{a_0+\sum_{l=1}^{t}\alpha^{-l}\exp\left(\boldsymbol{\beta}'\boldsymbol{x}_{il}\right)}. \qquad (3)
\end{aligned}$$

This formula is somewat similar to credibility estimators with geometric weights, as those considered, e.g., by Sundt (1988).

Note that whereas the total number of claims recorded in the past is a sufficient statistic of the claim history when reevaluating the claim frequency with static random effects, this is no more true with dynamic heterogeneity. Formula (3) shows that all the claim counts $N_{i1}, N_{i2}, \ldots, N_{it}$ have to be kept by the company to compute the a posteriori claim frequency.

The value of α can be estimated on the basis of historical data, for instance choosing $\hat{\alpha}$ in order to maximize the log-likelihood function of the Negative Binomial model. In order to write down this likelihood, we introduce

$$\Theta_{it}^{+} = \Theta_{it}\exp(\boldsymbol{\beta}'\boldsymbol{x}_{it}),$$

where

$$[\Theta_{it}^{+}|\mathcal{F}_{i,t-1}] \sim \mathcal{G}am(a_{i,t|t-1}^{+}, \tau_{i,t|t-1}^{+}),$$

with parameters $a_{i,t|t-1}^{+}$ and $\tau_{i,t|t-1}^{+}$ given by

$$\begin{aligned}
a_{i,t|t-1}^{+} &= a_{i,t|t-1} = \alpha a_{it-1},\\
\tau_{i,t|t-1}^{+} &= \exp\left(-\boldsymbol{\beta}'\boldsymbol{x}_{it}\right)\tau_{i,t|t-1}\\
&= \exp\left(-\boldsymbol{\beta}'\boldsymbol{x}_{it}\right)\alpha\tau_{it-1}.
\end{aligned}$$

The log-likelihood to be maximized with respect to α is then given by

$$\begin{aligned}
\log L(\alpha) = \sum_{i=1}^{n}\sum_{t=\xi+1}^{T_i}\Big(&\log\left(\Gamma\left(a_{it}\right)\right) - \log\left(\Gamma\left(N_{it}+1\right)\right)\\
&- \log\left(\Gamma\left(\alpha a_{it-1}\right)\right)\\
&+ \alpha a_{it-1}\log\left(\alpha\tau_{it-1}\exp\left(-\boldsymbol{\beta}'\boldsymbol{x}_{it}\right)\right)\\
&- a_{it}\log\left(\alpha\tau_{it-1}\exp\left(-\boldsymbol{\beta}'\boldsymbol{x}_{it}\right)+1\right)\Big),
\end{aligned}$$

where ξ is the index of the first non-zero observation. The value of α controls the rate of convergence to the minimal premium for claim-free policyholders. Of course, this value can also be tuned inspired by commercial considerations.

4 Numerical illustrations

In this section, we compare the experience rating model built in Section 3 to the classical approach pioneered by Dionne & Vanasse (1989, 1992) who extended classical credibility models to segmented tariffs. In their framework, the residual heterogeneity for policyholder i is represented by a random variable $\Theta_i \sim \mathcal{G}am(a,a)$. Given $\Theta_i = \theta$, the N_{it}'s are independent random variables, such that

$$N_{it} \sim \mathcal{P}oi\big(\exp(\boldsymbol{\beta}'\boldsymbol{x}_{it})\theta\big). \qquad (4)$$

The optimal estimator of the claim frequency for year $t+1$ given the past experience is

$$\mathbb{E}[N_{i,t+1}|\mathcal{F}_{it}] = \exp(\boldsymbol{\beta}'\boldsymbol{x}_{i,t+1})\frac{a+\sum_{\ell=0}^{t-1}N_{i,t-\ell}}{a+\sum_{\ell=0}^{t-1}\exp(\boldsymbol{\beta}'\boldsymbol{x}_{i,t-\ell})}. \qquad (5)$$

In the Negative Binomial model with static heterogeneity, the claims history enters the a posteriori claim frequency as an unweighted average of observed past claims counts, whereas (3) discounts old claims and thus recognizes that recent claim counts are more predictive than older ones.

It is worth mentioning that (5) is a particular case of (3) for $\alpha = 1$ and $a_{i0} \equiv a$. Note that only *a priori* tarification is used in the first period. Moreover, when the regression component is limited to a constant (1, say), one finds the usual Bayes estimator of the expected number of accidents caused by policyholder i during the $(t+1)$th coverage period.

4.1 The data

The working sample is the same used in Bolancé, Pinquet & Guillén (2003). It contains 80 994 policyholders, which represent 10% of the portfolio of a major Spanish insurance company. We selected only policies covering cars for private use and retained a balanced panel data set containing information since 1991 until 1997. Hence, all individual histories have the same duration. Table 1 contains the frequency of claims at fault from the first to the seventh period, as well as the observed claims distributions by year.

Period	Mean Frequency	0	1	2	3	4	5
1991	0.079	75 053	5 485	419	34	3	0
1992	0.070	75 716	4 910	351	16	1	0
1993	0.063	76 238	4 410	313	28	4	1
1994	0.064	76 267	4 340	356	30	1	0
1995	0.066	76 093	4 472	387	39	3	0
1996	0.069	75 847	4 751	354	33	9	0
1997	0.075	75 428	5 127	405	31	3	0
Total	0.069	530 642	33 495	2 585	211	24	1

Table 1. Claim distributions per calendar year.

We have twelve exogeneous variables that are kept in the panel plus the yearly number of accidents. The explanatory variables are detailed in Table 2. 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. Even if all the policies are observed for 7 years, some of their characteristics vary. For instance, the driver (whose gender, age, etc. are recorded in the databasis) may be replaced with another one, a new vehicle can be bought (more powerful than the preceding one), the policyholder may move from one zone to another, etc. Table 3 shows the relative frequencies of each regressor per period.

Variable	Definition
X_1	equals 1 for female drivers and 0 for males
X_2	equals 1 when driving zone is in urban area, 0 otherwise
X_3	equals 1 when driving zone is medium risk (Madrid and Catalonia)
X_4	equals 1 when driving zone is high risk (Northern Spain)
X_5	equals 1 if the driving license is between 4 and 14 years old
X_6	equals 1 if the driving license is 15 or more years old
X_7	equals 1 if the policyholder has been in the company between 3 and 5 years
X_8	equals 1 if the policyholder is in the company for more than 5 years
X_9	equals 1 if the driver is 30 years old or younger
X_{10}	equals 1 if coverage includes comprehensive (material damage only, fire excepted)
X_{11}	equals 1 if coverage includes comprehensive (material damage garantee and fire)
X_{12}	equals 1 if the power of the vehicle is larger than or equal to 5500cc

Table 2. Description of the explanatory variables included in the a priori risk evaluation.

	Period							
	1	2	3	4	5	6	7	Total
$\Pr[\widehat{X_1} = 1]$	0.14	0.14	0.15	0.15	0.16	0.16	0.17	0.15
$\Pr[\widehat{X_2} = 1]$	0.71	0.71	0.71	0.70	0.68	0.67	0.66	0.69
$\Pr[\widehat{X_3} = 1]$	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24
$\Pr[\widehat{X_4} = 1]$	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
$\Pr[\widehat{X_5} = 1]$	0.44	0.40	0.36	0.32	0.28	0.26	0.23	0.33
$\Pr[\widehat{X_6} = 1]$	0.52	0.57	0.61	0.66	0.70	0.72	0.74	0.64
$\Pr[\widehat{X_7} = 1]$	0.37	0.36	0.31	0.36	0.26	0.14	0.00	0.26
$\Pr[\widehat{X_8} = 1]$	0.27	0.37	0.54	0.64	0.74	0.86	1.00	0.63
$\Pr[\widehat{X_9} = 1]$	0.22	0.19	0.16	0.13	0.11	0.09	0.08	0.14
$\Pr[\widehat{X_{10}} = 1]$	0.17	0.17	0.16	0.16	015	0.16	0.16	0.16
$\Pr[\widehat{X_{11}} = 1]$	0.22	0.24	0.27	0.30	0.33	0.35	0.37	0.30
$\Pr[\widehat{X_{12}} = 1]$	0.71	0.74	0.76	0.78	0.79	0.81	0.82	0.77

Table 3. Descriptive statistics of rating factors per period.

4.2 The results

We proceed as follows. First, we use the observations relating to 1991 to estimate the regression coefficients $\beta_0, \beta_1, \ldots, \beta_{12}$ involved in the models (1)-(4), as well as the dispersion parameter. To this end, we use the maximum likelihood method in the Negative Binomial distribution. Table 4 presents the estimated regression coefficients and dispersion parameter in the Negative Binomial model using the 1991 data. Then, the annual expected claim frequencies are reevaluated through (3)-(5) given the number of claims recorded in the databasis.

Table 5 displays summary statistics about the updating coefficients in the model proposed by Dionne & Vanasse (1989,1992), that is, the so-called relativities involved in (5) and applied to the policyholders from 1992 to 1997. For each period the following measures are calculated. The average relativity is presented in the first column. Note that in the first period, i.e. year 1991, the relativity is equal to 100%. The following columns in Table 5 show the standard deviation, the first quartile, the median, the third quartile, and the 90% percentile. The average relativity decreases in the first periods and then slightly increases. The median decreases with the periods.

Parameter	Estimate
β_0	-2.2262
β_1	0.0125
β_2	-0.0645
β_3	-0.0983
β_4	0.2224
β_5	-0.3297
β_6	-0.3732
β_7	-0.1357
β_8	-0.2632
β_9	0.1546
β_{10}	0.2472
β_{11}	0.0681
β_{12}	0.0835
Dispersion	0.9899

Table 4. Parameter estimates in the Negative Binomial model for the 1991 data.

t	Mean	STD	Q1 (25%)	Median (P50%)	Q3 (75%)	P90%
1991	1.000	-	-	-	-	-
1992	1.010	2.090	0.421	0.459	0.504	0.535
1993	0.940	1.887	0.047	0.056	0.875	3.521
1994	0.937	1.897	0.045	0.053	0.829	3.723
1995	0.937	1.888	0.042	0.050	0.789	3.696
1996	0.939	1.862	0.041	0.048	0.752	3.590
1997	0.943	1.825	0.039	0.046	1.036	3.505
1998	0.951	1.775	0.037	0.045	1.124	3.442

Table 5. Relativities in the model of Dionne & Vanasse (1989,1992).

In Table 6 we present summary statistics about the average relativities obtained with the procedure using the Bayesian experience rating with time dependent random effects, i.e. those involved in (3). The maximum likelihood estimation of the parameter α is 0.1. In the columns we see the first quartile, the median, the third quartile, and the 90% percentile for each yearly period, in the sample portfolio.

t	Mean	STD	Q1 (25%)	Median (P50%)	Q3 (75%)	P90%
1991	1.000	-	-	-	-	-
1992	0.913	3.789	0.038	0.040	0.041	0.041
1993	0.984	3.913	0.001	0.002	0.002	0.598
1994	0.945	3.783	<0.001	<0.001	<0.001	0.612
1995	0.952	3.751	<0.001	<0.001	<0.001	0.596
1996	0.976	3.747	<0.001	<0.001	<0.001	0.589
1997	1.034	3.795	<0.001	<0.001	<0.001	0.610
1998	0.991	3.583	<0.001	<0.001	<0.001	0.603

Table 6. Relativities in the experience rating model with dynamic random effects and $\alpha = 0.1$.

The results obtained for the Bayesian experience rating with time dependent random effects show a very marked bonus-hunger effect compared to those derived in the model proposed by Dionne & Vanasse (1989,1992). Since the estimation of the α parameter is closer to zero than to one, every policyholder is very much influenced by his/her own experience. So the higher relativities are suffered by those few who had claims at fault. The average and median relativities are very small. After some yearly periods drivers with no accidents pay a very moderate amount of premium. The conclusion is that dynamic random effects produce premiums much more disperse than those derived with a static random effect.

5 Conclusion

The results obtained in this paper show that the Bayesian method with dynamic heterogeneity provides a neat separation of the policyholders. The relativies are so extreme that it would not be possible to implement them in practice, because a small subset of policyholders would pay extremely high premia while others would pay very small premia. The information acquired after a claim induces a strong effect on the expected number of claims for the subsequent years. The effect is magnified by the weighting parameter α. The approach presented here is useful for identifying those policyholders with a bad claim record. It should not be recommended to price the automobile insurance contract if the principle of risk sharing (pooling) is to be perceived by consumers.

The α parameter can be used to compare the claiming experience in different portfolios, where a larger α would mean that there is less dynamicity in the unobserved risk factors.

The prediction in the Harvey-Fernandes model has the same weakness as in Sundt (1988): assuming that the total credibility converges towards unity as time goes on, an exponential decay of the credibility of periods as a function of seniority leads to overweight recent information.

Acknowledgements

The authors wish to thank the anonymous referee for useful comments, which improved the original manuscript.

Michel Denuit acknowledges the financial support of the *Communauté française de Belgique* under contract "Projet d'Actions de Recherche Concertées" ARC 04/09-320, as well as the financial support of the *Banque Nationale de Belgique* under grant "Risk measures and Economic capital".

REFERENCES

[1] Albrecht, P. (1985). An evolutionary credibility model for claim numbers. ASTIN Bulletin 15, 1-17.

[2] Bolancé, C., Pinquet, J., & Guillén, M. (2003). Time-varying credibility for frequency risk models. Insurance: Mathematics and Economics 33, 272-282.

[3] Brannas, K., & Hellstrom, J. (2001). Generalizations to the AR(1) count data model. Econometric Reviews 20, 425-443.

[4] Dionne, G., & Vanasse, C. (1989). A generalization of actuarial automobile insurance rating models: the Negative Binomial distribution with a regression component. ASTIN Bulletin 19, 199-212.

[5] Dionne, G., & Vanasse, C. (1992). Automobile insurance ratemaking in the presence of asymmetrical information. Journal of Applied Econometrics 7, 149-165.

[6] Freeland, R.K., & McCabe, B.P.M. (2004). Forecasting discrete valued low count time series. International Journal of Forecasting 20, 427-434.

[7] Gerber, H.U., & Jones, D.A. (1975). Cedibility formulas of the updating type. In "Credibility", edited by Kahn, P.M., Academic Press, New York, pp. 89-105.

[8] Gouriéroux, C., & Jasiak, J. (2004). Heterogeneous INAR(1) model with application to car insurance. Insurance: Mathematics & Economics 34, 177-192.

[9] Harvey, A.C., and Fernandes, C. (1989). Time series models for count or qualitative observations. Journal of Business & Economics Statistics 7, 407-422 (with discussion).

[10] Jewell, W.S. (1975). Model variations in credibility theory. In "Credibility", edited by Kahn, P.M., Academic Press, New York, pp. 193-244.

[11] Jewell, W.S. (1976). Two classes of covariance matrices giving simple linear forecasts. Scandinavian Actuarial Journal, 15-29.

[12] Lambert, Ph. (1996a). Modelling of repeated series of count data measured at unequally spaced times. Applied Statistics 45, 31-38.

[13] Lambert, Ph. (1996b). Modeling of nonlinear growth curve on series of correlated count data measured at unequally spaced times: a full likelihood based approach. Biometrics 52, 50-55.

[14] Pinquet, J., Guillén, M., & Bolancé, C. (2001). Allowance for the age of claims in bonus-malus systems. Astin Bulletin 31, 337-348.

[15] Purcaru, O., Guillén, M., & Denuit, M. (2004). Linear credibility models based on time series for claim counts. Belgian Actuarial Bulletin 4, 62-74.

[16] Sundt, B. (1981). Recursive credibility estimation. Scandinavian Actuarial Journal, 3-21.

[17] Sundt, B. (1983). Finite credibility formula in evolutionary model. Scandinavian Actuarial Journal, 106-116.

[18] Sundt, B. (1988). Credibility estimators with geometric weights. Insurance: Mathematics & Economics 7, 113-122.

[19] Yau, K.K.W., Lee, A.H., & Carrivick, P.J.W. (2004). Modelling zero-inflated count series with application to occupational health. Computer Methods and Programs in Biomedicine 74, 47-52.