Sarmanov Family of Multivariate Distributions for Bivariate Dynamic Claim Counts Model

Anas Abdallah*, Jean-Philippe Boucher†, and Hélène Cossette‡

November 26, 2015

Abstract

To predict future claims, it is well-known that the most recent claims are more predictive than older ones. However, classic panel data models for claim counts, such as the multivariate negative binomial distribution, do not put any time weight on past claims. More complex models can be used to consider this property, but often need numerical procedures to estimate parameters. When we want to add a dependence between different claim count types, the task would be even more difficult to handle. In this paper, we propose a bivariate dynamic model for claim counts, where past claims experience of a given claim type is used to better predict the other type of claims. This new bivariate dynamic distribution for claim counts is based on random effects that come from the Sarmanov family of multivariate distributions. To obtain a proper dynamic distribution based on this kind of bivariate priors, an approximation of the posterior distribution of the random effects is proposed. The resulting model can be seen as an extension of the dynamic heterogeneity model described in Bolancé et al. (2007). We apply this model to two samples of data from a major Canadian insurance company, where we show that the proposed model is one of the best models to adjust the data. We also show that the proposed model allows more flexibility in computing predictive premiums because closed-form expressions can be easily derived for the predictive distribution, the moments and the predictive moments.

Keywords: Poisson-Gamma mixture, Sarmanov, Maximum Likelihood Estimation, a Posteriori Ratemaking, Dynamic Claim Count.

*École d'Actuariat, Université Laval
†Département de mathématiques, UQAM
‡École d'Actuariat, Université Laval

1 Introduction

One of the most critical problems in property and casualty insurance is to determine future numbers of claims and cost of claims. A related task is the calculation of the premium, i.e ratemaking. Parametric modeling of these random variables allows to identify the risk level through explanatory variables and clarifies the behavior of insureds. In this paper, we will focus only on the frequency part of this task, i.e. the modeling of the number of claims.

Risk classification techniques for claim counts have been the topic of many papers in the actuarial literature. For example, Denuit et al. (2007) provide an exhaustive overview of count data models for insurance claims. In recent years, dependence between all the contracts of the same insured has been supposed in actuarial models, leading to what is called panel data modeling. Panel data modeling allows the premiums to depend on past claims experience, where the classic credibility theory can be used. In this paper, panel data models for claim counts are generalized in two ways: 1) by allowing different claim types to be modeled simultaneously, and 2) by allowing a time weight for past claims, because we expect that the most recent claims are more predictive than the oldest ones. To our knowledge, the proposed model is the first parametric model with continuous random effects to achieve these generalizations.

In actuarial sciences, the modeling of two different types of claims has already been studied. For example, Pinquet (1998) uses Poisson residuals to create dependence between at-fault and not-at-fault claims in automobile insurance, while Boucher and Inoussa (2014) use Bonus-Malus Systems with specific penalty rules allowing different claim types to affect the premium. Frees and Valdez (2008) also model various type of claims by decomposing all possibilities of claim types that may occur for a single accident. Generalizations to time-dependent heterogeneous models have also often been studied in the actuarial literature. To obtain a dynamic approach with continuous random effects, a parametric model would normally need T-dimensional integrals to express the joint distribution of all claims of a single insured (Xu et al. (2007)). Consequently, complex numerical procedures that are not suited for panel data framework are sometimes needed (see for example Jung and Liesenfeld (2001)). Other approaches have been proposed to put a dynamic effect into count models: evolutionary credibility models in Gerber and Jones (1975), Jewell (1975), Poisson residuals in Pinquet et al. (2001), or more recently copulas with the jittering method in Shi and Valdez (2014).

In this paper, to obtain this generalization of panel data models, the bivariate claim count distribution will be based on two conditional Poisson distributions with two gamma random effects distributions. Dependence will be supposed between the random effects, based on the Sarmanov family of multivariate distributions. This family of multivariate distributions has nice properties. Indeed, we show that this family of distributions offers flexibility in the choice of marginals and allows a closed-form expression for the joint density function. Additionally, we show that the posterior density of the bivariate random effects has approximately the same form as its prior. In particular, we show that the proposed model allows closed-form expressions for the predictive distribution, and a closed-form expression for the predictive premium, which can be an important insight for the insurer. Note that even if the illustrated model is used based on Poisson-gamma combinations, the proposed model can be easily used to generalize models with different conditional distributions or different random effects distributions.

In Section 2, we review the modeling of claim counts, where notations are set and random effects defined. In Section 3, we define the Sarmanov family of multivariate distribution. A multivariate extension of the dynamic model based on Harvey and Fernandes (1989) is presented in Section 4. The proposed model can be seen as an extension of the dynamic heterogeneity model described in Bolancé et al. (2007). To be able to use such a dynamic approach, an approximation of the *a posteriori* Sarmanov distribution of the random effects is proposed, where it is supposed that this *a posteriori* distribution has the same form as the *a priori* distribution. Using data from a sample of a major Canadian insurance company, two numerical illustrations are performed in Section 5, where different claim types are used. Predictive premiums as well as the predictive variance are also computed and compared for various models. Section 6 concludes the paper.

2 Claim Count Modeling

2.1 General notations

We are interested in modeling the number of claims $N_{i,\ell,t}$, for the ith policyholder ($i = 1, ..., n$) of an insurance portfolio, of a given type of claim ℓ ($\ell = 1, 2$), at time t ($t = 1, ..., T$). To simplify the notations, subscript i will be removed for the remainder of the paper. To construct our model, we will suppose a conditional Poisson distribution of mean $\lambda_{\ell,t}\theta_\ell$, i.e.

$$(N_{\ell,t} \mid \Theta_\ell = \theta_\ell) \;\sim\; \text{Poisson}\left(\lambda_{\ell,t}\theta_\ell\right),$$

where $\lambda_{\ell,t} = \exp(\beta' x_{\ell,t})$ and $x_{\ell,t}$ represents the vector of all the pertinent covariates for claim type ℓ during year t. Classically, most of the ratemaking techniques rely on generalized linear models (GLM) (see McCullagh and Nelder (1989)) to estimate the regression parameters.

For each claim type, hidden characteristics are usually captured by an additional random term that affects all the contracts of the same insured. Each random effect is denoted by the random variable $\Theta_\ell, \ell = 1, 2$. Even if each claim type shares common hidden characteristics, we will first suppose that Θ_1 and Θ_2 are independent. This assumption will be relaxed later.

We assume that each Θ_ℓ is gamma distributed with shape parameter α_ℓ and scale parameter τ_ℓ. Both parameters α_ℓ and τ_ℓ are first considered stationary. Hence, we have

$$\Theta_\ell \;\sim\; \text{Gamma}\left(\alpha_\ell, \tau_\ell\right),$$

with probability density function (pdf) denoted by

$$h(\theta_\ell;\alpha_\ell,\tau_\ell)=\frac{\tau_\ell^{\alpha_\ell}}{\Gamma(\alpha_\ell)}\theta_\ell^{\alpha_\ell-1}\exp(-\tau_\ell\theta_\ell).$$

Let us denote by $\mathbf{N}_\ell=(N_{\ell,1},...,N_{\ell,T})$ the vector of the number of claims, and $f_{N_{\ell,t}}(n_{\ell,t}\mid\Theta_\ell=\theta_\ell)$ the discrete conditional probability mass function of $(N_{\ell,t}\mid\Theta_\ell=\theta_\ell)$. Consequently, the joint probability mass function (pmf) of $\mathbf{N}_\ell$, denoted by $f_{\mathbf{N}_\ell}(\boldsymbol{n}_\ell;\alpha_\ell,\tau_\ell)$, is given by

$$\begin{aligned}
f_{\mathbf{N}_\ell}(\boldsymbol{n}_\ell;\alpha_\ell,\tau_\ell) &= \Pr(\mathbf{N}_\ell=\mathbf{n}_\ell)\\
&= \int_0^\infty f_{\mathbf{N}_\ell}(\mathbf{n}_\ell|\Theta_\ell=\theta_\ell)h(\theta_\ell;\tau_\ell,\alpha_\ell)d\theta_\ell\\
&= \left(\prod_{t=1}^{T}\frac{\lambda_{\ell,t}^{n_{\ell,t}}}{n_{\ell,t}!}\right)\frac{\Gamma(n_{\ell,\bullet}+\alpha_\ell)}{\Gamma(\alpha_\ell)}\left(\frac{\tau_\ell}{\lambda_{\ell,\bullet}+\tau_\ell}\right)^{\alpha_\ell}(\lambda_{\ell,\bullet}+\tau_\ell)^{-n_{\ell,\bullet}}, \qquad (2.1)
\end{aligned}$$

which corresponds to the joint pdf of a multivariate negative binomial random vector (MVNB), with $n_{\ell,\bullet}=\sum_{t=1}^{T}n_{\ell,t}$ and $\lambda_{\ell,\bullet}=\sum_{t=1}^{T}\lambda_{\ell,t}$. See Boucher et al. (2008) for details. In the stationary case, for parameter identification, we suppose that $\alpha_\ell=\tau_\ell$. In this case, the marginal moments of $N_{\ell,t}$ are given by

$$E[N_{\ell,t}]=\lambda_{\ell,t}\frac{\alpha_\ell}{\tau_\ell}=\lambda_{\ell,t}\quad\text{and}\quad \mathrm{Var}(N_{\ell,t})=\lambda_{\ell,t}\frac{\alpha_\ell}{\tau_\ell}+\lambda_{\ell,t}^2\frac{\alpha_\ell}{\tau_\ell^2}=\lambda_{\ell,t}+\frac{\lambda_{\ell,t}^2}{\alpha_\ell}.$$

For ratemaking purposes, $E[N_{\ell,t}]$ is often called the *a priori* premium because it is the premium charged to new insureds, or insureds without claims experience.

2.2 Predictive Distribution

The random effect term models the heterogeneity of the model and incorporates the hidden characteristics. Consequently, it is reasonable to believe that these hidden characteristics are partly revealed by the number of claims reported by the policyholders. Indeed, at each insured period, the random effects can be updated given the past claim experience, revealing some insured-specific information. Henceforth, insightful information can be retrieved from the claim experience.

The Poisson and gamma distributions are natural conjugates, thus the *a posteriori* distribution is again a gamma distribution with updated parameters $\alpha_\ell^*=\alpha_\ell+\sum_{t=1}^{T}n_{\ell,t}$ and $\tau_\ell^*=\tau_\ell+\sum_{t=1}^{T}\lambda_{\ell,t}$ (see Boucher et al. (2008) for details). Thus, the predictive mean can be expressed as

$$E[N_{\ell,T+1}\mid N_{\ell,1},...,.N_{\ell,T}] = \lambda_{\ell,T+1}\frac{\alpha_\ell^*}{\tau_\ell^*}=\lambda_{\ell,T+1}\frac{\alpha_\ell+\sum_{t=1}^{T}n_{\ell,t}}{\alpha_\ell+\sum_{t=1}^{T}\lambda_{\ell,t}}. \qquad (2.2)$$

In a ratemaking context, this is often called the predictive premium. We clearly observe how past experience is incorporated in the computation of the predictive mean. Indeed, we can see that all past claims have equal weight in the predictive premium calculation, meaning that an old claim increases the premium as much as a newer claim does. A more intuitive model would suppose that both Θ_ℓ, $\ell = 1, 2$, can evolve over time, resulting in a model where the most recent claims are more predictive than the oldest ones.

3 Sarmanov Family of Bivariate Distributions

Sarmanov's bivariate distribution was introduced in the literature by Sarmanov (1966), and was also proposed in physics by Cohen (1984) under a more general form. Lee (1996) suggests a multivariate version and discusses several applications in medicine. Recently, due to its flexible structure, Sarmanov's bivariate distribution gained interest in different applied studies. For example, Schweidel et al. (2008) use a bivariate Sarmanov model to capture the relationship between a prospective customer's time until acquisition of a particular service and the subsequent duration for which the service is retained. Miravete (2009) presents two models based on the Sarmanov distribution and uses them to compare the number of tariff plans offered by two competing cellular telephone companies. Danaher and Smith (2011) discuss applications to marketing (see also the references therein). In the insurance field, Hernández-Bastida and Fernández-Sánchez (2012) use the bivariate Sarmanov distribution for premium evaluation, more recently Abdallah et al. (2015) use this family of distributions to show its suitability in a loss reserving context. In this paper, we use the Sarmanov distribution to accommodate correlation of unknown characteristics of a driver that might impact all types of claims simultaneously.

3.1 Definitions

Let the random couple $\mathbf{\Theta} = (\Theta_1, \Theta_2)$ have a bivariate Sarmanov distribution, with gamma marginals

$$u^S(\theta_1, \theta_2) = h(\theta_1; \alpha_1, \tau_1)\, h(\theta_2; \alpha_2, \tau_2)\, (1 + \omega \phi_1(\theta_1)\, \phi_2(\theta_2)), \tag{3.1}$$

where ϕ_ℓ, $\ell = 1, 2$ are two bounded non-constant functions such that $\int_{-\infty}^{\infty} \phi_\ell(t)\, u_\ell(t)\, dt = 0$ and ω is a real number that satisfies the condition

$$1 + \omega \phi_1(\theta_1)\, \phi_2(\theta_2) \geq 0 \text{ for all } \theta_\ell\ ,\ i \in \{1, 2\}.$$

One of the main interesting properties of the Sarmanov distribution is that the multivariate distribution can support a wide range of marginals, such as the gamma distribution. Different methods are proposed in Lee (1996) to construct mixing functions ϕ_ℓ for different types of marginals. As mentioned in Lee (1996), different types of mixing functions can be used to yield different multivariate distributions with the same set of marginals. Based on Corollary 2 in Lee (1996), a mixing function can be defined as $\phi_\ell(\theta_\ell) = \exp(-\theta_\ell) - L_\ell$, where L_ℓ is the Laplace transform of the marginal distribution evaluated at 1. Hence, given our choice of distribution for Θ_ℓ, $\ell = 1, 2$, we have

$$\phi_\ell(\theta_\ell) = \exp(-\theta_\ell) - \left(\frac{\tau_\ell}{1+\tau_\ell}\right)^{\alpha_\ell}.$$

As for the dependence parameter ω of the bivariate Sarmanov distribution, in the case of gamma marginals, it is bounded as follows $B_{\text{inf}} < \omega < B_{\text{sup}}$ with

$$B_{\text{inf}} = \frac{-1}{\max\left\{\left(\frac{\tau_1}{1+\tau_1}\right)^{\alpha_1}\left(\frac{\tau_2}{1+\tau_2}\right)^{\alpha_2}, \left(1-\left(\frac{\tau_1}{1+\tau_1}\right)^{\alpha_1}\right)\left(1-\left(\frac{\tau_2}{1+\tau_2}\right)^{\alpha_2}\right)\right\}}$$

$$B_{\text{sup}} = \frac{1}{\max\left\{\left(\frac{\tau_1}{1+\tau_1}\right)^{\alpha_1}\left(1-\left(\frac{\tau_2}{1+\tau_2}\right)^{\alpha_2}\right), \left(\frac{\tau_2}{1+\tau_2}\right)^{\alpha_2}\left(1-\left(\frac{\tau_1}{1+\tau_1}\right)^{\alpha_1}\right)\right\}}.$$

This result is given in Corollary 2 of Lee (1996). Consequently, for gamma marginals, and using the notations of the previous section, the prior joint pdf of (Θ_1, Θ_2) is given by

$$\begin{aligned} u^S(\theta_1, \theta_2) &= (1+\vartheta)\, h(\theta_1; \alpha_1, \tau_1) h(\theta_2; \alpha_2, \tau_2) + \vartheta h(\theta_1; \alpha_1, \tau_1+1) h(\theta_2; \alpha_2, \tau_2+1) \\ &\quad -\vartheta h(\theta_1; \alpha_1, \tau_1) h(\theta_2; \alpha_2, \tau_2+1) - \vartheta h(\theta_1; \alpha_1, \tau_1+1) h(\theta_2; \alpha_2, \tau_2), \end{aligned} \tag{3.2}$$

where $\vartheta = \omega\left(\frac{\tau_1}{1+\tau_1}\right)^{\alpha_1}\left(\frac{\tau_2}{1+\tau_2}\right)^{\alpha_2}$. This last expression corresponds to a linear combination of the product of univariate (gamma) pdf's and highlights the attractive features of the Sarmanov family of distributions.

3.2 Bivariate Count Distributions

A critical problem when modeling dependence between claim counts is to obtain a closed-form expression for the joint distribution. The Sarmanov distribution will be a good ally to circumvent this problem. Let us denote by $f_{\mathbf{N}_1,\mathbf{N}_2}$ the discrete joint probability mass function of $(\mathbf{N}_1, \mathbf{N}_2)$, i.e. $f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1, \mathbf{n}_2) = \Pr(\mathbf{N}_1 = \mathbf{n}_1, \mathbf{N}_2 = \mathbf{n}_2)$ which can be expressed as

$$\begin{aligned} f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1, \mathbf{n}_2) &= \int_0^\infty \int_0^\infty f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1, \mathbf{n}_2 | \Theta_1 = \theta_1, \Theta_2 = \theta_2) u^S(\theta_1, \theta_2) d\theta_1 d\theta_2 \\ &= (1+\vartheta)\, f_{\mathbf{N}_1}(\boldsymbol{n}_1; \alpha_1, \tau_1) f_{\mathbf{N}_2}(\boldsymbol{n}_2; \alpha_2, \tau_2) + \vartheta f_{\mathbf{N}_1}(\boldsymbol{n}_1; \alpha_1, \tau_1+1) f_{\mathbf{N}_2}(\boldsymbol{N}_2; \alpha_2, \tau_2+1) \\ &\quad -\vartheta f_{\mathbf{N}_1}(\boldsymbol{N}_1; \alpha_1, \tau_1) f_{\mathbf{N}_2}(\boldsymbol{N}_2; \alpha_2, \tau_2+1) - \vartheta f_{\mathbf{N}_1}(\boldsymbol{N}_1; \alpha_1, \tau_1+1) f_{\mathbf{N}_2}(\boldsymbol{N}_2; \alpha_2, \tau_2). \end{aligned} \tag{3.3}$$

Note that we obtain a linear combination of products of MVNB distributions. The simplicity and form of the model greatly facilitate many calculations, such as the moments of the distribution. For the *a priori* mean, we obtain the following result:

$$E\left[N_{1,t}+N_{2,t}\right]=\lambda_{1,t}\frac{\alpha_1}{\tau_1}+\lambda_{2,t}\frac{\alpha_2}{\tau_2}, \tag{3.4}$$

Note that the mean of the sum is the same as the one obtained for the sum of two MVNB distributions. However, a covariance term is added to the sum of the variance, as shown in the following result:

$$\begin{aligned}\mathrm{Var}(N_{1,t}+N_{2,t}) &= \lambda_{1,t}\frac{\alpha_1}{\tau_1}+\lambda_{1,t}^2\frac{\alpha_1}{\tau_1^2}+\lambda_{2,t}\frac{\alpha_2}{\tau_2}+\lambda_{2,t}^2\frac{\alpha_2}{\tau_2^2}\\ &+2\lambda_{1,t}\lambda_{2,t}\left(\vartheta\frac{\alpha_1}{\tau_1}\frac{\alpha_2}{\tau_2}+\vartheta\frac{\alpha_1}{\tau_1+1}\frac{\alpha_2}{\tau_2+1}-\vartheta\frac{\alpha_1}{\tau_1+1}\frac{\alpha_2}{\tau_2}-\vartheta\frac{\alpha_1}{\tau_1}\frac{\alpha_2}{\tau_2+1}\right).\end{aligned} \tag{3.5}$$

Similarly to what Purcaru and Denuit (2002) did for univariate claim count models, it would be interesting to analyze the dependence induced by this kind of model. Recently, Bolancé et al. (2014) show that the Sarmanov family of distributions has upper tail dependence equal to zero when the marginal distributions have tail Gumbel type, as Gamma distribution. However, this impact is mitigated in our context and findings, because the marginals of the Sarmanov family of multivariate distributions represent the heterogeneity components in our model. The interpretation of such a finding in our context would mean that the probability that the couple (Θ_1,Θ_2) has two extreme heterogeneity components tends to zero.

3.2.1 Predictive Joint Distribution

As done with the univariate analysis of Section 2.2, the posterior distribution is also useful as it reveals insured-specific information. Because of the bivariate structure of the random effects, past claims experience of a given claim type can be used to update the random effects distribution of the other type of claims. The posterior bivariate joint density function of the couple (Θ_1,Θ_2) conditioned on $(\mathbf{N}_1,\mathbf{N}_2)$ is given by

$$\begin{aligned}u^S(\theta_1,\theta_2\mid\mathbf{n}_1,\mathbf{n}_2) &= \psi_1 h(\theta_1;\alpha_1^*,\tau_1^*)h(\theta_2;\alpha_2^*,\tau_2^*)+\psi_2 h(\theta_1;\alpha_1^*,\tau_1^*+1)h(\theta_2;\alpha_2^*,\tau_2^*+1)\\ &\quad -\psi_3 h(\theta_1;\alpha_1^*,\tau_1^*)h(\theta_2;\alpha_2^*,\tau_2^*+1)-\psi_4 h(\theta_1;\alpha_1^*,\tau_1^*+1)h(\theta_2;\alpha_2^*,\tau_2^*),\end{aligned} \tag{3.6}$$

where $\alpha_\ell^*=\alpha_\ell+n_{\ell,\bullet}$, $\tau_\ell^*=\tau_\ell+\lambda_{\ell,\bullet}$, $\ell=1,2$, and

$$
\begin{aligned}
\psi_1 &= \frac{1}{f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1,\mathbf{n}_2)}\left(1+\vartheta\right) f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_1,\tau_1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_2,\tau_2)\\
\psi_2 &= \frac{1}{f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1,\mathbf{n}_2)}\vartheta f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_1,\tau_1+1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_2,\tau_2+1)\\
\psi_3 &= \frac{1}{f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1,\mathbf{n}_2)}\vartheta f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_1,\tau_1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_2,\tau_2+1)\\
\psi_4 &= \frac{1}{f_{\mathbf{N}_1,\mathbf{N}_2}(\mathbf{n}_1,\mathbf{n}_2)}\vartheta f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_1,\tau_1+1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_2,\tau_2).
\end{aligned}
$$

This last expression shows that the posterior bivariate density function of (Θ_1, Θ_2), is again a linear combination of the product of univariate gamma pdfs. The posterior density is hence called a pseudo-conjugate to the prior density (Lee (1996)) in the sense that the posterior density is a linear combination of products of densities from the univariate natural exponential family of distributions (gamma in our case).

The joint predictive distribution of $\mathbf{N}_1$ and $\mathbf{N}_2$ at time $T+1$, given all the past observations up to time T can also be computed. This will enable us to evaluate notably the expected annual claim frequency conditionally on past experience. The Sarmanov distribution allows us to obtain a closed-form expression for this joint prediction. Indeed, using (2.1) and (3.6), we get

$$
\begin{aligned}
&f_{N_{1,T+1},N_{2,T+1}|\mathbf{N}_1,\mathbf{N}_2}\left(n_{1,T+1},n_{2,T+1}\right) =\\
&\quad \psi_1 f_{N_{1,T+1}|\mathbf{N}_1}\left(n_{1,T+1};\alpha_1+n_{1,\bullet},\tau_1+\lambda_{1,\bullet}\right) f_{N_{2,T+1}|\mathbf{N}_2}\left(n_{2,T+1};\alpha_2+n_{2,\bullet},\tau_2+\lambda_{2,\bullet}\right)\\
&\qquad +\psi_2 f_{N_{1,T+1}|\mathbf{N}_1}\left(n_{1,T+1};\alpha_1+n_{1,\bullet},\tau_1+1+\lambda_{1,\bullet}\right) f_{N_{2,T+1}|\mathbf{N}_2}\left(n_{2,t+1};\alpha_2+n_{2,\bullet},\tau_2+1+\lambda_{2,\bullet}\right)\\
&\qquad -\psi_3 f_{N_{1,T+1}|\mathbf{N}_1}\left(n_{1,T+1};\alpha_1+n_{1,\bullet},\tau_1+\lambda_{1,\bullet}\right) f_{N_{2,T+1}|\mathbf{N}_2}\left(n_{2,T+1};\alpha_2+n_{2,\bullet},\tau_2+1+\lambda_{2,\bullet}\right)\\
&\qquad -\psi_4 f_{N_{1,T+1}|\mathbf{N}_1}\left(n_{1,T+1};\alpha_1+n_{1,\bullet},\tau_1+1+\lambda_{1,\bullet}\right) f_{N_{2,T+1}|\mathbf{N}_2}\left(n_{2,T+1};\alpha_2+n_{2,\bullet},\tau_2+\lambda_{2,\bullet}\right),
\end{aligned}
\tag{3.7}
$$

with ψ_j, $j=1,2,3,4$ as given in (3.6) and $n_{\ell,\bullet}$ and $\lambda_{\ell,\bullet}$ as given in (2.1), where we suppose

$$
\begin{aligned}
f_{N_{\ell,T+1}|\mathbf{N}_\ell}\left(n_{\ell,t+1};\alpha_*,\tau_*\right) &= \left(\frac{\lambda_{\ell,t+1}^{n_{\ell,t+1}}}{n_{\ell,t+1}!}\right)\frac{\Gamma\left(n_{\ell,t+1}+\alpha_*\right)}{\Gamma(\alpha_*)}\frac{\tau_*^{\alpha_*}}{(\lambda_{\ell,t+1}+\tau_*)^{n_{\ell,t+1}+\alpha_*}}\\
&= \frac{\Gamma(\alpha_*+n_{\ell,t+1})}{\Gamma(\alpha_*)\Gamma(n_{\ell,t+1}+1)}\left(\frac{\tau_*}{\lambda_{\ell,t+1}+\tau_*}\right)^{\alpha_*}\left(\frac{\lambda_{\ell,t+1}}{\lambda_{\ell,t+1}+\tau_*}\right)^{n_{\ell,t+1}},
\end{aligned}
\tag{3.8}
$$

which corresponds to a negative binomial distribution with parameters α_* and $\dfrac{\tau_*}{\lambda_{\ell,t+1}+\tau_*}$.

One of the main advantages of using the Sarmanov family of bivariate distributions is the possibility to derive closed-form expressions for the mean and variance of the total future number of claims. Let $\Im_{\ell,T}$ be the history of claim counts of type ℓ up to time T. Mathematically, $\Im_{\ell,T}$ is the sigma algebra generated by the random variables $N_{\ell,1}, N_{\ell,2}, ..., N_{\ell,T}$, with $\Im_T = (\Im_{1,T}, \Im_{2,T})$ and $N^{\text{tot}}_{T+1} = N_{1,T+1} + N_{2,T+1}$. It can be shown that the total expected annual claim frequency for year $T+1$ is

$$E\left[N^{\text{tot}}_{T+1} \mid \Im_T\right] = \lambda_{1,T+1}\left((\psi_1-\psi_4)\frac{\alpha_1^*}{\tau_1^*} + (\psi_2-\psi_3)\frac{\alpha_1^*}{\tau_1^*+1}\right) + \lambda_{2,T+1}\left((\psi_1-\psi_3)\frac{\alpha_2^*}{\tau_2^*} + (\psi_2-\psi_4)\frac{\alpha_2^*}{\tau_2^*+1}\right), \tag{3.9}$$

with $\alpha_\ell^* = \alpha_\ell + n_{\ell,\bullet}$, $\tau_\ell^* = \tau_\ell + \lambda_{\ell,\bullet}$, $\ell = 1,2$. The total variance of the sum can be expressed as

$$\begin{aligned}
\text{Var}(N^{\text{tot}}_{T+1} \mid \Im_T) &= \lambda_{1,T+1}\left((\psi_1-\psi_4)\frac{\alpha_1^*}{\tau_1^*} + (\psi_2-\psi_3)\frac{\alpha_1^*}{\tau_1^*+1}\right) \\
&+ \lambda_{1,T+1}^2\left((\psi_1-\psi_4)\frac{\alpha_1^*}{\tau_1^{*2}} + (\psi_2-\psi_3)\frac{\alpha_1^*}{(\tau_1^*+1)^2}\right) \\
&+ \lambda_{2,T+1}\left((\psi_1-\psi_3)\frac{\alpha_2^*}{\tau_2^*} + (\psi_2-\psi_4)\frac{\alpha_2^*}{\tau_2^*+1}\right) \\
&+ \lambda_{2,T+1}^2\left((\psi_1-\psi_3)\frac{\alpha_2^*}{\tau_2^{*2}} + (\psi_2-\psi_4)\frac{\alpha_2^*}{(\tau_2^*+1)^2}\right) \\
&+ 2\lambda_{1,T+1}\lambda_{2,T+1}\left(\psi_1\frac{\alpha_1^*}{\tau_1^*}\frac{\alpha_2^*}{\tau_2^*} + \psi_2\frac{\alpha_1^*}{\tau_1^*+1}\frac{\alpha_2^*}{\tau_2^*+1} - \psi_3\frac{\alpha_1^*}{\tau_1^*+1}\frac{\alpha_2^*}{\tau_2^*} - \psi_4\frac{\alpha_1^*}{\tau_1^*}\frac{\alpha_2^*}{\tau_2^*+1}\right) \\
&- 2\left(\lambda_{1,T+1}\left((\psi_1-\psi_4)\frac{\alpha_1^*}{\tau_1^*} + (\psi_2-\psi_3)\frac{\alpha_1^*}{\tau_1^*+1}\right)\right)\left(\lambda_{2,t+1}\left((\psi_1-\psi_3)\frac{\alpha_2^*}{\tau_2^*} + (\psi_2-\psi_4)\frac{\alpha_2^*}{\tau_2^*+1}\right)\right),
\end{aligned} \tag{3.10}$$

with $\alpha_\ell^* = \alpha_\ell + n_{\ell,\bullet}$, $\tau_\ell^* = \tau_\ell + \lambda_{\ell,\bullet}$, $\ell = 1,2$.

It is worth mentioning that, as expected, the model borrows past information from one series to predict future claim counts of the other series. Indeed, the terms ψ_j for $j = 1,2,3,4$ depend on $N_{\ell,1}, N_{\ell,2}, ..., N_{\ell,T}$, $\ell = 1,2$, and are used to compute the expected value of each type of claim. The dependence parameter ω intervenes in the predictive mean and variance computations through the same terms ψ_j. Simplified and developed expressions of the terms ψ_j's are presented later in the paper in another context.

4 Multivariate Dynamic Random Effects

4.1 Model and motivations

Models where the random effects Θ_ℓ evolve over time would normally need T-dimensional integrals to express the joint distribution of all claims of a single insured. Consequently, complex numerical procedures or approximated inference methods are sometimes needed (see for example Jung and Liesenfeld (2001) or Xu et al. (2007)). Other approaches have been proposed to put a dynamic effect into count models: evolutionary credibility models in

Gerber and Jones (1975), Jewell (1975), Poisson residuals in Pinquet et al. (2001), or more recently copulas with the jittering method in Shi and Valdez (2014).

In our paper, we propose an extension of Bolancé et al. (2007), which is based on the idea of Harvey and Fernandes (1989). We note this model as the H-F model, referring directly to Harvey-Fernandes. Their model supposes that the risk characteristics are captured through a dynamic effect $\Theta_{\ell,t}$, which is considered evolutionary and time-dependent, i.e. that its distribution evolves over time and is updated through past experience. Formally, the classic Poisson-gamma model described in Section 2 is generalized and allows the underlying risk parameter to vary in successive periods, with the following dynamic:

$$(\Theta_{\ell,t} \mid \Im_{\ell,t}) \sim \text{Gamma}\left(\alpha_{\ell,t}, \tau_{\ell,t}\right), \ \ell \in \{1,2\}. \tag{4.1}$$

It is also supposed that

$$(\Theta_{\ell,t} \mid \Im_{\ell,t-1}) \sim \text{Gamma}\left(\alpha_{\ell,t|t-1}, \tau_{\ell,t|t-1}\right), \ \ell \in \{1,2\}, \tag{4.2}$$

where

$$\begin{cases} \alpha_{\ell,t|t-1} = \nu_\ell \alpha_{\ell,t-1} \\ \tau_{\ell,t|t-1} = \nu_\ell \tau_{\ell,t-1}. \end{cases} \tag{4.3}$$

The parameter ν_ℓ is a weighting parameter less than or equal to 1. The initial conditions of the dynamic model, i.e. the distribution of $\Theta_{\ell,1}$, is supposed $\text{Gamma}\left(\alpha_{\ell,0}, \tau_{\ell,0}\right)$, with $\alpha_{\ell,0} = \tau_{\ell,0}$. This means that the premium for the first year equals $\lambda_{\ell,1}$, because $E\left[\Theta_{\ell,1}\right] = 1$.

Using Bayes' theorem, the posterior distribution for $(\Theta_{\ell,t} \mid \Im_{\ell,t})$ is again a gamma distribution with updated parameters

$$\begin{cases} \alpha_{\ell,t} = \nu_\ell \alpha_{\ell,t-1} + n_{\ell,t} \\ \tau_{\ell,t} = \nu_\ell \tau_{\ell,t-1} + \lambda_{\ell,t}. \end{cases}$$

By induction, the above parameters can be expressed recursively as follows:

$$\begin{cases} \alpha_{\ell,t} = \left(\nu_\ell\right)^t \alpha_{\ell,0} + \sum_{k=0}^{t-1} \left(\nu_\ell\right)^k n_{\ell,t-k} \\ \tau_{\ell,t} = \left(\nu_\ell\right)^t \alpha_{\ell,0} + \sum_{k=0}^{t-1} \left(\nu_\ell\right)^k \lambda_{\ell,t-k}. \end{cases} \tag{4.4}$$

Given the past experience, the resulting joint distribution of $\mathbf{N}_\ell = N_{\ell,1}, ..., N_{\ell,T}$ can then be expressed as

$$f_{\mathbf{N}_\ell}(\mathbf{n}_\ell; \alpha_{\ell,t}, \tau_{\ell,t}) = \left(\prod_{t=1}^{T} \frac{\lambda_{\ell,t}^{n_{\ell,t}}}{n_{\ell,t}!}\right) \frac{\Gamma(n_{\ell,\bullet} + \alpha_{\ell,t|t-1})}{\Gamma(\alpha_{\ell,t|t-1})} \left(\frac{\tau_{\ell,t|t-1}}{\lambda_{\ell,\bullet} + \tau_{\ell,t|t-1}}\right)^{\alpha_{\ell,t|t-1}} \left(\lambda_{\ell,\bullet} + \tau_{\ell,t|t-1}\right)^{-n_{\ell,\bullet}}, \tag{4.5}$$

where $\alpha_{\ell,t|t-1}$ and $\tau_{\ell,t|t-1}$ are as given in (4.3).

We observe that the multivariate joint distribution is similar to equation (2.1), but the random effects parameters are now time-dependent. Note that unlike the stationary model, where the sum of claim counts was a sufficient statistic, the dynamic model keeps the time period of each claim. The *a priori* moments of the H-F model, are given by

$$E[N_{\ell,t}] = \lambda_{\ell,t} \ \text{ and } \ \text{Var}(N_{\ell,t}) = \lambda_{\ell,t} + \frac{\lambda_{\ell,t}^2}{\alpha_{\ell,0}},$$

meanwhile the predictive moments can be expressed as

$$E[N_{\ell,T+1} \mid N_{\ell,1}, ..., .N_{\ell,T}] = \lambda_{\ell,T+1}\frac{\alpha_{\ell,T}}{\tau_{\ell,T}} \ \text{ and } \ \text{Var}(N_{\ell,T+1}) = \lambda_{\ell,T+1}\frac{\alpha_{\ell,T}}{\tau_{\ell,T}} + \lambda_{\ell,T+1}^2\frac{\alpha_{\ell,T}}{\tau_{\ell,T}^2},$$

where $\alpha_{\ell,T}$ and $\tau_{\ell,T}$ are obtained following equation (4.4) with $\ell = 1, 2$. We observe that the *a priori* and predictive moments of this model have the same expressions as in the MVNB model, with time-dependent underlying parameters.

4.2 Sarmanov distribution and dynamic heterogeneity

One of the main advantages of the Sarmanov family of bivariate distributions is its pseudo-conjugate property for the posterior distribution (see equation (3.6)). However, this property might not be sufficient to directly suppose a dynamic structure for the Sarmanov distribution. Indeed, to be able to assume a dynamic approach with the Sarmanov distribution, like the one proposed for the Poisson-gamma model (or MVNB) in equation (4.4), the bivariate *a posteriori* distribution of the random effects needs to be a conjugate to the prior, where updated parameters α_ℓ^*, τ_ℓ^*, $\ell = 1, 2$ can be modified easily using a structure similar to the equations above. To obtain such a structure, the *a posteriori* distribution of the correlated random effects needs to be, once again, a member of the family of Sarmanov multivariate distribution. As just specified, the Sarmanov family of bivariate distributions does not possess this conjugate property, but its pseudo-conjugate property might enable us to construct an interesting alternative.

The posterior distribution of random effects obtained in (3.6) is a weighted sum of posterior gamma distributions, with $\alpha_\ell^* = \alpha_\ell + n_{\ell,\bullet}$, $\tau_\ell^* = \tau_\ell + \lambda_{\ell,\bullet}$, $\ell = 1, 2$. The difference between what would have been called a conjugate distribution and the pseudo-conjugate comes from $\psi_1, \psi_2, \psi_3, \psi_4$ that are not expressed solely in terms of α_ℓ^* and τ_ℓ^*, but also α_ℓ and τ_ℓ. We propose to modify the posterior distribution to obtain a distribution that is only a function of α_ℓ^* and τ_ℓ^*. This modification will be the first step to obtain a dynamic bivariate count distribution. The proposed posterior Sarmanov distribution, now referred to as Approximated Sarmanov, is then expressed as:

$$\begin{aligned} u^{S^*}(\theta_1, \theta_2 \mid \mathbf{n}_1, \mathbf{n}_2) \ = \ & (1 + \vartheta^*)\, h(\theta_1; \alpha_1^*, \tau_1^*) h(\theta_2; \alpha_2^*, \tau_2^*) + \vartheta^* h(\theta_1; \alpha_1^*, \tau_1^* + 1) h(\theta_2; \alpha_2^*, \tau_2^* + 1) \\ & - \vartheta^* h(\theta_1; \alpha_1^*, \tau_1^*) h(\theta_2; \alpha_2^*, \tau_2^* + 1) - \vartheta^* h(\theta_1; \alpha_1^*, \tau_1^* + 1) h(\theta_2; \alpha_2^*, \tau_2^*), \end{aligned} \tag{4.6}$$

where the distribution has the same form as the *a priori* distribution of (Θ_1, Θ_2) with updated parameters $\vartheta^* = \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*}$, where $\alpha_\ell^* = \alpha_\ell + n_{\ell,\bullet}$, $\tau_\ell^* = \tau_\ell + \lambda_{\ell,\bullet}$, for $\ell = 1, 2$.

4.2.1 Quality of the Approximation

The Approximated Sarmanov distribution for random effects has the desired properties to be generalized into a dynamic approach. However, before adding the dynamic structure, we need to quantify the approximation of the *a posteriori* Sarmanov distribution. The difference between the Sarmanov and Approximated Sarmanov can be expressed as:

$$\begin{aligned}
&u^{S^*}(\theta_1, \theta_2 \mid \mathbf{n}_1, \mathbf{n}_2) - u^S(\theta_1, \theta_2 \mid \mathbf{n}_1, \mathbf{n}_2) = \\
&\quad \delta_1 h(\theta_1; \alpha_1^*, \tau_1^*) h(\theta_2; \alpha_2^*, \tau_2^*) + \delta_2 h(\theta_1; \alpha_1^*, \tau_1^* + 1) h(\theta_2; \alpha_2^*, \tau_2^* + 1) \\
&\quad - \delta_3 h(\theta_1; \alpha_1^*, \tau_1^*) h(\theta_2; \alpha_2^*, \tau_2^* + 1) - \delta_4 h(\theta_1; \alpha_1^*, \tau_1^* + 1) h(\theta_2; \alpha_2^*, \tau_2^*).
\end{aligned} \tag{4.7}$$

Each term δ_j, $j = 1, 2, 3, 4$ can be simplified as:

$$\begin{aligned}
\delta_1 &= \left[1 + \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*}\right] - \psi_1 \\
&= \vartheta^* \left(1 - \frac{L_1/L_1^* + L_2/L_2^* - 1}{1 + \vartheta + \vartheta^*(1 - L_1/L_1^* - L_2/L_2^*)}\right) \\
\delta_2 &= \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*} - \psi_2 \\
&= \vartheta^* \left(1 - \frac{1}{1 + \vartheta + \vartheta^*\left(1 - L_1/L_1^* - L_2/L_2^*\right)}\right) \\
\delta_3 &= \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*} - \psi_3 \\
&= \vartheta^* \left(1 - \frac{L_2/L_2^*}{1 + \vartheta + \vartheta^*\left(1 - L_1/L_1^* - L_2/L_2^*\right)}\right) \\
\delta_4 &= \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*} - \psi_4. \\
&= \vartheta^* \left(1 - \frac{L_1/L_1^*}{1 + \vartheta + \vartheta^*\left(1 - L_2/L_2^* - L_1/L_1^*\right)}\right),
\end{aligned}$$

where $\vartheta^* = \omega \left(\frac{\tau_1^*}{1+\tau_1^*}\right)^{\alpha_1^*} \left(\frac{\tau_2^*}{1+\tau_2^*}\right)^{\alpha_2^*}$, $L_\ell = \left(\frac{\tau_\ell}{1+\tau_\ell}\right)^{\alpha_\ell}$, and $L_\ell^* = \left(\frac{\tau_\ell^*}{1+\tau_\ell^*}\right)^{\alpha_\ell^*}$, for $\ell = 1, 2$. It is worth-mentioning that the differences δ_j's are complementary and offsetting one another, resulting in a sum of differences equal to zero. This condition comes from the fact that (4.7) is a difference between two proper distributions.

We analyzed the approximation for different values of the parameters. The Approximated Sarmanov is identical to the Sarmanov distribution when $\omega = 0$, which is the particular case of independent random effects. We also observed that the values of δ_j, for $j = 1, ..., 4$ are proportional to ω. The difference caused by the approximation also depends on the parameters α_ℓ^*, τ_ℓ^*, and thus depends on the time of each claim and also on sums of past claims, i.e. $n_{\ell,\bullet}$ for $\ell = 1, 2$. In Figure 4.1, for a specific choice of parameters, as a function of $n_{\ell,\bullet}$ for $\ell = 1, 2$, we illustrate the values of the δ_j's for a large observation period $T = 100$.

We see that the approximation is less accurate in the cases where $n_{1,\bullet}$ and $n_{2,\bullet}$ tend to behave inversely. This represents unusual situations because claims of different types are assumed to be positively correlated, which means that an insured with a high value of $n_{1,\bullet}$ should also normally have a high number of past claims of type 2. When the numbers of past claims are similar for each type, the approximation seems to be more accurate and reasonable.

We also analyze the model for smaller time periods because it is more realistic for insurance data. Indeed, insurance datasets are usually constructed with $T < 10$ (see Boucher et al. (2008) for example). We illustrate the values of the δ_j's for $T = 5$ in Figure 4.2. Our analysis shows that the differences expressed by the δ_j's are closer to zero (note that the scales are different than those of Figure 4.1). Henceforth, one observes differences between the Approximated Sarmanov and the original Sarmanov models, but the highest differences occur for unusual situations. However, empirically, most insureds are located in the area $n_{j,t} = \{0, 1\}$, for $j = 1, 2$ where the differences are much smaller. Thus, for small values of T, the Approximated Sarmanov distribution is close to the original Sarmanov, but it is important to understand that they are not the same model.

4.2.2 Dynamic Sarmanov

The closed-form expressions for the moments of the Approximated Sarmanov do not change from those obtained in (3.5). It can be shown that the Approximated Sarmanov model generates a predictive annual claim frequency given by

$$\begin{aligned} E\left[N_{T+1}^{\text{tot}} \mid \Im_T\right] &= \left((1+\vartheta^*-\vartheta^*)\,\frac{\alpha_1^*}{\tau_1^*} + (\vartheta^*-\vartheta^*)\,\frac{\alpha_1^*}{\tau_1^*+1}\right) + \lambda_{2,T+1}\left((1+\vartheta^*-\vartheta^*)\,\frac{\alpha_2^*}{\tau_2^*} + (\vartheta^*-\vartheta^*)\,\frac{\alpha_2^*}{\tau_2^*+1}\right), \\ &= \lambda_{1,T+1}\frac{\alpha_1^*}{\tau_1^*} + \lambda_{2,t+1}\frac{\alpha_2^*}{\tau_2^*} \\ &= \lambda_{1,T+1}\frac{\alpha_1 + \sum_{t=1}^T n_{1,t}}{\tau_1 + \sum_{t=1}^T \lambda_{1,t}} + \lambda_{2,T+1}\frac{\alpha_2 + \sum_{t=1}^T n_{2,t}}{\tau_2 + \sum_{t=1}^T \lambda_{2,t}}, \end{aligned} \tag{4.8}$$

where α_ℓ^* and τ_ℓ^*, $\ell = 1, 2$, are given by (2.1). The predictive variance is expressed as follows

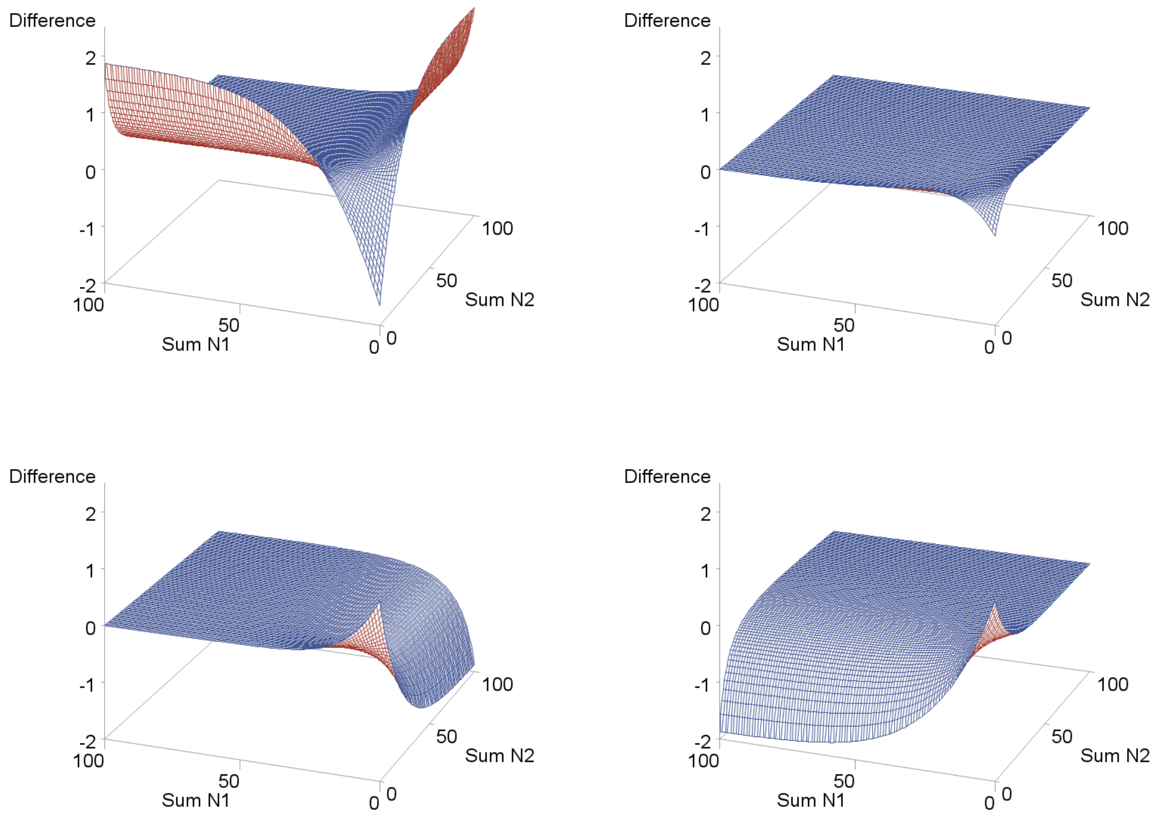

Figure 4.1: Graphs of δ_1, δ_2, δ_3 and δ_4 respectively, for a time series of $T = 100$ periods, with $\lambda_1 = \lambda_2 = 0.15$, $\alpha_1 = \alpha_2 = 0.7$ and $\omega = 2$

$$
\begin{aligned}
\operatorname{Var}(N_{T+1}^{\text{tot}} \mid \Im_T) &= \lambda_{1,T+1}\frac{\alpha_1^*}{\tau_1^*} + \lambda_{1,T+1}^2\frac{\alpha_1^*}{\tau_1^{*2}} + \lambda_{2,T+1}\frac{\alpha_2^*}{\tau_2^*} + \lambda_{2,T+1}^2\frac{\alpha_2^*}{\tau_2^{*2}} \\
&+ 2\lambda_{1,T+1}\lambda_{2,T+1}\vartheta^*\left(\frac{\alpha_1^*}{\tau_1^*}\frac{\alpha_2^*}{\tau_2^*} + \frac{\alpha_1^*}{\tau_1^*+1}\frac{\alpha_2^*}{\tau_2^*+1} - \frac{\alpha_1^*}{\tau_1^*+1}\frac{\alpha_2^*}{\tau_2^*} - \frac{\alpha_1^*}{\tau_1^*}\frac{\alpha_2^*}{\tau_2^*+1}\right),
\end{aligned}
\tag{4.9}
$$

where $\alpha_\ell^* = \alpha_\ell + n_{\ell,\bullet}$, $\tau_\ell^* = \tau_\ell + \lambda_{\ell,\bullet}$, $\ell = 1, 2$.

We observe that we obtain closed-from expressions for the predictive mean and variance, which is convenient for premium calculation. However, we can see that the model does not use the parameter ω in the calculation of the predictive mean. Moreover, the premium for insurance coverage ℓ only uses α_ℓ^* and τ_ℓ^*, which are based on information of the claim type ℓ only. Consequently, for a specific type of claim, the Approximated Sarmanov model cannot borrow information from the other types of claims in predictive modeling. This is contradic-

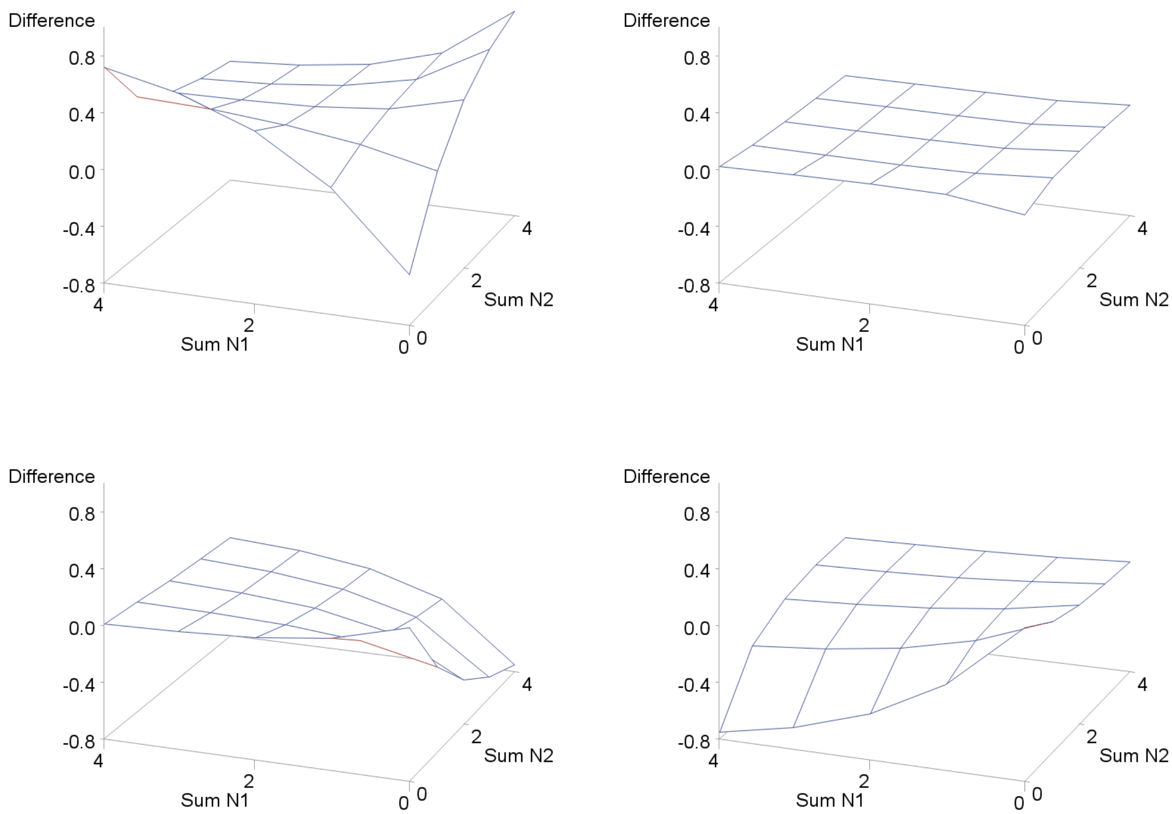

Figure 4.2: Graphs of δ_1, δ_2, δ_3 and δ_4 respectively, for a time series of $T = 5$ periods, with $\lambda_1 = \lambda_2 = 0.15$, $\alpha_1 = \alpha_2 = 0.7$ and $\omega = 2$

tory to the objective of our model because we expect that past claims experience of a given claim type should be used to better predict future claims of another correlated claim type.

One way to correct this gap in the model is to consider a dynamic model that extends the one proposed for the H-F model. Instead of adding only a weighting parameter ν_ℓ in the model, we also introduce other parameters γ_1 and γ_2 to borrow information from the claim types 2 and 1, respectively. We intuitively believe that this modification allows us to better predict future claims of each type. Formally, the parameters of such a model can be expressed as follows:

$$\begin{cases} \alpha_{1,t|t-1} = \nu_1 \left(\alpha_{1,t-1} + \gamma_1 n_{2,t} \right) \\ \tau_{1,t|t-1} = \nu_1 \left(\tau_{1,t-1} + \gamma_1 \lambda_{2,t} \right), \end{cases} \tag{4.10}$$

and

$$\begin{cases} \alpha_{2,t|t-1} = \nu_2 \left(\alpha_{2,t-1} + \gamma_2 n_{1,t}\right) \\ \tau_{2,t|t-1} = \nu_2 \left(\tau_{2,t-1} + \gamma_2 \lambda_{1,t}\right), \end{cases} \tag{4.11}$$

where ν_ℓ is again a weighting parameter less than 1. It can be shown by induction that the following general recursive relation holds:

$$\begin{cases} \alpha_{1,t} = \left(\nu_1\right)^t \alpha_{1,0} + \sum_{k=0}^{t-1} \left((\nu_1)^k n_{1,t-k} + (\nu_1)^{k+1} \gamma_1 n_{2,t-k} \right) \\ \tau_{1,t} = \left(\nu_1\right)^t \alpha_{1,0} + \sum_{k=0}^{t-1} \left((\nu_1)^k \lambda_{1,t-k} + (\nu_1)^{k+1} \gamma_1 \lambda_{2,t-k} \right), \end{cases} \tag{4.12}$$

and

$$\begin{cases} \alpha_{2,t} = \left(\nu_2\right)^t \alpha_{2,0} + \sum_{k=0}^{t-1} \left((\nu_2)^k n_{2,t-k} + (\nu_2)^{k+1} \gamma_2 n_{1,t-k} \right) \\ \tau_{2,t} = \left(\nu_2\right)^t \alpha_{2,0} + \sum_{k=0}^{t-1} \left((\nu_2)^k \lambda_{2,t-k} + (\nu_2)^{k+1} \gamma_2 \lambda_{1,t-k} \right). \end{cases} \tag{4.13}$$

With these proposed parameters, the *a posteriori* distribution and the predictive analysis incorporate information on the correlated type of claims and borrow insightful past experience of a given claim type to better predict the correlated claim type. As a generalization of the H-F model, the resulting model also puts time weight on the correlated claim. This is reflected by (4.12) and (4.13).

The proposed modification allows us to construct a dynamic structure for the bivariate count model with Sarmanov random effects. The *a posteriori* distribution of $(\Theta_{1,t}, \Theta_{2,t})$ for the Dynamic Sarmanov model is assumed to have the same form as the *a priori* joint pdf of $(\Theta_{1,t}, \Theta_{2,t})$, and can be expressed as

$$\begin{aligned} u^S\left(\theta_{1,t}, \theta_{2,t} \mid \Im_t\right) \;=\; & \left(1 + \omega \left(\frac{\tau_{1,t}}{1+\tau_{1,t}}\right)^{\alpha_{1,t}} \left(\frac{\tau_{2,t}}{1+\tau_{2,t}}\right)^{\alpha_{2,t}}\right) h(\theta_{1,t}; \alpha_{1,t}, \tau_{1,t}) h(\theta_{2,t}; \alpha_{2,t}, \tau_{2,t}) \\ & + \omega \left(\frac{\tau_{1,t}}{1+\tau_{1,t}}\right)^{\alpha_{1,t}} \left(\frac{\tau_{2,t}}{1+\tau_{2,t}}\right)^{\alpha_{2,t}} h(\theta_{1,t}; \alpha_{1,t}, \tau_{1,t}+1) h(\theta_{2,t}; \alpha_{2,t}, \tau_{2,t}+1) \\ & - \omega \left(\frac{\tau_{1,t}}{1+\tau_{1,t}}\right)^{\alpha_{1,t}} \left(\frac{\tau_{2,t}}{1+\tau_{2,t}}\right)^{\alpha_{2,t}} h(\theta_{1,t}; \alpha_{1,t}, \tau_{1,t}) h(\theta_{2,t}; \alpha_{2,t}, \tau_{2,t}+1) \\ & - \omega \left(\frac{\tau_{1,t}}{1+\tau_{1,t}}\right)^{\alpha_{1,t}} \left(\frac{\tau_{2,t}}{1+\tau_{2,t}}\right)^{\alpha_{2,t}} h(\theta_{1,t}; \alpha_{1,t}, \tau_{1,t}+1) h(\theta_{2,t}; \alpha_{2,t}, \tau_{2,t}), \end{aligned} \tag{4.14}$$

where $\alpha_{\ell,t}$ and $\tau_{\ell,t}$ are non-stationary parameters given by (4.12) and (4.13).

Note that when $\omega = 0$, the model can be seen as a bivariate version of the H-F model, noted Bivariate H-F. Hence, given known past experience, the joint distribution of $(\mathbf{N}_1, \mathbf{N}_2)$ for the Dynamic Sarmanov model has the following closed-form expression:

$$
\begin{aligned}
f_{\mathbf{N}_1,\mathbf{N}_2}(\boldsymbol{n}_1,\boldsymbol{n}_2) \;=\; & (1+\vartheta_t)\, f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_{1,t},\tau_{1,t}) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_{2,t},\tau_{2,t}) \\
& +\vartheta_t f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_{1,t},\tau_{1,t}+1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_{2,t},\tau_{2,t}+1) \\
& -\vartheta_t f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_{1,t},\tau_{1,t}) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_{2,t},\tau_{2,t}+1) \\
& -\vartheta_t f_{\mathbf{N}_1}(\boldsymbol{n}_1;\alpha_{1,t},\tau_{1,t}+1) f_{\mathbf{N}_2}(\boldsymbol{n}_2;\alpha_{2,t},\tau_{2,t}),
\end{aligned}
\tag{4.15}
$$

with $\alpha_{\ell,t}$ and $\tau_{\ell,t}$ given by (4.12) and (4.13) for $\ell = 1, 2$, where $\vartheta_t = \omega \left(\frac{\tau_{1,t}}{1+\tau_{1,t}}\right)^{\alpha_{1,t}} \left(\frac{\tau_{2,t}}{1+\tau_{2,t}}\right)^{\alpha_{2,t}}$. Note that the dependence parameter ω does not depend on time.

The moments of the model can be expressed in closed-form, and do not change from those obtained in equation (3.5). For the predictive premium and variance of the Dynamic Sarmanov model, it can be reduced to the following

$$
\begin{aligned}
E\left[N^{\text{tot}}_{T+1} \mid \Im_T\right] \;=\; & \lambda_{1,T+1}\frac{\alpha_{1,T}}{\tau_{1,T}} + \lambda_{2,T+1}\frac{\alpha_{2,T}}{\tau_{2,T}} \\
=\; & \lambda_{1,T+1}\frac{(\nu_1)^T \alpha_{1,0} + \sum_{k=0}^{T-1}\left(\nu_1^k n_{1,T-k} + \nu_1^{k+1}\gamma_1 n_{2,T-k}\right)}{(\nu_1)^T \alpha_{1,0} + \sum_{k=0}^{T-1}\left(\nu_1^k \lambda_{1,T-k} + \nu_1^{k+1}\gamma_1 \lambda_{2,T-k}\right)} \\
& +\lambda_{2,T+1}\frac{(\nu_2)^T \alpha_{2,0} + \sum_{k=0}^{T-1}\left(\nu_2^k n_{2,T-k} + \nu_2^{k+1}\gamma_2 n_{1,T-k}\right)}{(\nu_2)^T \alpha_{2,0} + \sum_{k=0}^{T-1}\left(\nu_2^k \lambda_{2,T-k} + \nu_2^{k+1}\gamma_2 \lambda_{1,T-k}\right)}
\end{aligned}
\tag{4.16}
$$

and

$$
\begin{aligned}
\text{Var}(N^{\text{tot}}_{T+1} \mid \Im_T) \;=\; & \lambda_{1,T+1}\frac{\alpha_{1,T}}{\tau_{1,T}} + \lambda_{1,T+1}^2\frac{\alpha_{1,T}}{\tau_{1,T}^2} + \lambda_{2,T+1}\frac{\alpha_{2,T}}{\tau_{2,T}} + \lambda_{2,T+1}^2\frac{\alpha_{2,T}}{\tau_{2,T}^2} \\
& + 2\lambda_{1,T+1}\lambda_{2,T+1}\vartheta_T\left(\frac{\alpha_{1,T}}{\tau_{1,T}}\frac{\alpha_{2,T}}{\tau_{2,T}} + \frac{\alpha_{1,T}}{\tau_{1,T}+1}\frac{\alpha_{2,T}}{\tau_{2,T}+1} - \frac{\alpha_{1,T}}{\tau_{1,T}+1}\frac{\alpha_{2,T}}{\tau_{2,T}} - \frac{\alpha_{1,T}}{\tau_{1,T}}\frac{\alpha_{2,T}}{\tau_{2,T}+1}\right),
\end{aligned}
\tag{4.17}
$$

where $\vartheta_T = \omega \left(\frac{\tau_{1,T}}{1+\tau_{1,T}}\right)^{\alpha_{1,T}} \left(\frac{\tau_{2,T}}{1+\tau_{2,T}}\right)^{\alpha_{2,T}}$, $\alpha_{\ell,T}$ and $\tau_{\ell,T}$ are again obtained from equations (4.12) and (4.13), with $\ell = 1, 2$.

We observe that the mean of a given claim type uses the past information of the correlated type of claims, through the crossed parameters γ_1 and γ_2. This link between the claim types does not directly depend on the dependence parameter ω, as it was the case for the stationary Sarmanov model. However, the ω intervenes in the calculation of the predictive variance of the Dynamic Sarmanov model, which can be a crucial additional information for various premium principles.

Variable	Description
X1	equals 1 if the insured is between 16 and 25 years old
X2	equals 1 if the insured is between 26 and 60 years old
X3	equals 1 if the vehicle is 0 years old
X4	equals 1 if the vehicle is 1-3 years old
X5	equals 1 if the vehicle is 4-5 years old
X6	equals 1 if the insured owns a home
X7	equals 1 if there is only one driver
X8	equals 1 if there are two drivers
X9	equals 1 if the insured is single
X10	equals 1 if the insured is divorced
X11	equals 1 if the insured has no minor convictions

Table 5.1: Binary variables summarizing the information available about each policyholder.

5 Empirical Illustration

5.1 Data used

We implement all the models presented in this paper with a sample of insurance data that comes from a major Canadian insurance company. Only private used cars have been considered in this sample. We consider 11 exogenous variables, shown in Table 5.1. For every policy we have the initial information at the beginning of the period to describe the profile of the driver. The unbalanced panel data contain information from 2003 to 2008. The sample contains 79,755 insurance contracts, which come from 26,251 policyholders.

The empirical illustration is performed on two pairs of claims types: collision vs comprehensive (noted pair COL/COM) and at-fault vs non-at-fault collision claims (noted pair AF/NAF). We decided to work with two different empirical illustrations to better describe the behavior of our models. We thus expose the models to a wider possibility of situations, which allows us to better analyze their performance and better highlight their properties.

Comprehensive coverage protects damage to the car that results from covered perils not related to a collision. Namely, a scenario that could cause damage to the car that has nothing to do with striking another vehicle. In many cases, this can include theft, vandalism, fire, natural disasters like a hurricane or a tornado, falling objects, etc. Thus, one would expect that if the accident is really a pure comprehensive accident, it should not give any indication of the competence of the driver or better predict future collisions. However, dependence may come from unobservable risk characteristics. In fact, some insureds tend to claim more than others, regardless of the type of claim. This might be explained by a social context as well, in the sense that an insured who lives in a riskier area could be exposed to both types of claims. Moreover, this dependence might also be caused by several factors, such as the driving competence of a driver (collision with a vehicle and collision with an object are often positively correlated), but this might also be explained by the behavior of the insured. Hence, the use of a model that allows dependence between coverages is justified.

In the second illustration, the collision coverage is separated into at-fault and non-at-fault claims. If the non-at-fault claims were really defined as pure bad luck, meaning that they have nothing to do with the behavior of the insured, then it would be irrational to believe that non-at-fault claims would be correlated with at-fault claims. However, in Canada, non-at-fault claims correspond to specific type of accidents, more related to the car's location in the accident. This is well known in Canada, and even if insurers cannot increase the premium for non-at-fault claims, insurers must sometimes find original ways to penalize drivers with non-at-fault claims (see Boucher and Inoussa (2014)). Consequently, for possibly the same reasons cited above for collision and comprehensive coverages that might lead to dependence, it seems logical to believe that dependence can exist between these two types of collision claims as well.

5.2 Model Calibration

Tables 5.2 and 5.3 exhibit the fit statistics along with the estimated parameters for the MVNB distribution, compared to the most popular count distributions, i.e. the Poisson and Negative Binomial type-2 (NB2) distributions. When the null hypothesis is on the boundary of the parameter space, a correction must be done to the likelihood ratio test, namely one-sided statistic tests (see Boucher et al. (2007) for more details). Consequently, a modified likelihood ratio test has been used to check if the Poisson is rejected against the NB2 or against the MVNB for both datasets. On the other hand, because NB2 and MVNB are non-nested models, the Akaike Information Criterion (AIC) has to be preferred to compare the models. In our case, despite the fact that the Poisson cannot always be rejected against the NB2 for all four coverages, we observe that the MVNB model is preferred over Poisson and NB2 for all coverages. This conclusion is interesting because unlike the Poisson or the NB2 distributions, the MVNB distribution allows temporal dependence between all contracts of the insured. This means that our intuition that a premium should somewhat depend on past claim experience is confirmed.

We also fit the Sarmanov and the Approximated Sarmanov models for comparison purposes, to validate the quality of the approximation supposed in the construction of the Approximated Sarmanov. Results are shown in Tables 5.4 and 5.5. We calculated the loglikelihood on the Approximated Sarmanov by using the estimated parameters of the Sarmanov model. We obtain close loglikelihood values ($-25,531.98$ vs. $-25,532.55$) for the pair COL/COM, meaning that models are close. Additionally, expressed with 2 decimals, we observe that the optimized loglikelihood of the Approximated Sarmanov ($-25,532.55$) is approximately the same as the one calculated with the MLE parameters of the Sarmanov distribution. For the pair AF/NAF, the loglikelihood obtained by Approximated Sarmanov by using the MLE of the Sarmanov distribution is equal to $-20,884.63$, while the maximum loglikelihood obtained by the Sarmanov is equal to $-20,878.76$, a slightly higher difference. The maximum loglikelihood obtained with the Approximated Sarmanov is also a little bit

	Poisson				NB2				MVNB			
	COL		COM		COL		COM		COL		COM	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-2.5517	0.0847	-4.1896	0.1641	-2.5474	0.0913	-4.1889	0.1604	-2.5739	0.0904	-4.1912	0.1709
β_{X1}	0.6349	0.0717	0.4553	0.1388	0.6419	0.0769	0.4562	0.1409	0.6416	0.0774	0.4644	0.1468
β_{X2}	0.2231	0.0470	0.3962	0.0854	0.2229	0.0501	0.3962	0.0858	0.2261	0.0508	0.3990	0.0877
β_{X3}	0.3369	0.0413	0.3927	0.0695	0.3383	0.0436	0.3924	0.0711	0.3207	0.0434	0.3885	0.0719
β_{X4}	0.3298	0.0382	0.3944	0.0643	0.3314	0.0402	0.3948	0.0650	0.3228	0.0400	0.3924	0.0660
β_{X5}	0.2245	0.0408	0.0919	0.0721	0.2250	0.0427	0.0923	0.0738	0.2149	0.0423	0.0907	0.0749
β_{X6}	0.0904	0.0310	0.1392	0.0528	0.0917	0.0331	0.1395	0.0542	0.0929	0.0328	0.1400	0.0541
β_{X7}	-0.4855	0.0564	-0.0731	0.1191	-0.4905	0.0607	-0.0736	0.1198	-0.4780	0.0615	-0.0775	0.1232
β_{X8}	-0.3888	0.0577	-0.0781	0.1195	-0.3917	0.0623	-0.0781	0.1252	-0.3891	0.0629	-0.0830	0.1258
β_{X9}	0.1901	0.0375	0.0940	0.0643	0.1927	0.0393	0.0943	0.0647	0.1932	0.0401	0.0942	0.0663
β_{X10}	0.2097	0.0566	0.2080	0.0929	0.2134	0.0595	0.2085	0.0947	0.2031	0.0611	0.2126	0.0955
β_{X11}	-0.2349	0.0490	-0.1445	0.0884	-0.2356	0.0525	-0.1448	0.0928	-0.2054	0.0529	-0.1377	0.0886
α_ℓ					1.3170	0.1331	0.6615	0.3875	0.6855	0.0651	0.6769	0.1722
LogLik	-25,635.6				-25,542.26				-25,532.79			
AIC	51,319.20				51,136.52				51,117.58			

Table 5.2: Parameter estimation - Stationary models for the pair COL/COM

	Poisson				NB2				MVNB			
	AF		NAF		AF		NAF		AF		NAF	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-3.3021	0.1258	-3.1918	0.1162	-3.3037	0.1282	-3.1893	0.1173	-3.3132	0.1374	-3.1967	0.1200
β_{X1}	0.7298	0.1031	0.5407	0.1003	0.7310	0.1045	0.5409	0.1016	0.7360	0.1103	0.5412	0.1040
β_{X2}	0.1586	0.0705	0.2745	0.0636	0.1586	0.0707	0.2738	0.0648	0.1606	0.0758	0.2747	0.0668
β_{X3}	0.3745	0.0630	0.3085	0.0549	0.3752	0.0633	0.3090	0.0555	0.3696	0.0648	0.3035	0.0560
β_{X4}	0.3960	0.0582	0.2786	0.0513	0.3963	0.0580	0.2792	0.0515	0.3949	0.0594	0.2760	0.0519
β_{X5}	0.3097	0.0608	0.1571	0.0547	0.3104	0.0611	0.1576	0.0554	0.3067	0.0628	0.1544	0.0554
β_{X6}	0.1230	0.0468	0.0648	0.0420	0.1230	0.0473	0.0642	0.0422	0.1242	0.0482	0.0654	0.0428
β_{X7}	-0.5720	0.0811	-0.4153	0.0754	-0.5717	0.0845	-0.4141	0.0766	-0.5726	0.0869	-0.4115	0.0798
β_{X8}	-0.4568	0.0833	-0.3329	0.0769	-0.4567	0.0857	-0.3325	0.0786	-0.4587	0.0880	-0.3322	0.0815
β_{X9}	0.2097	0.0558	0.1749	0.0497	0.2094	0.0567	0.1745	0.0505	0.2130	0.0584	0.1742	0.0511
β_{X10}	0.1536	0.0885	0.2501	0.0737	0.1536	0.0888	0.2502	0.0746	0.1545	0.0896	0.2461	0.0759
β_{X11}	-0.2442	0.0726	-0.2273	0.0659	-0.2433	0.0730	-0.2298	0.0663	-0.2295	0.0759	-0.2208	0.0694
α_i					0.1112	0.1620	0.2950	0.1678	0.4943	0.1330	0.3478	0.0939
LogLik	-20,919.09				-20,917.05				-20,899.84			
AIC	41,886.18				41,886.10				41,851.68			

Table 5.3: Parameter estimation - Stationary models for the pair AF/NAF

	Sarmanov				Approximated Sarmanov			
	COL		COM		COL		COM	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-2.5739	0.0923	-4.1927	0.1655	-2.5736	0.0885	-4.1905	0.1686
β_{X1}	0.6423	0.0773	0.4645	0.1395	0.6415	0.0770	0.4646	0.1434
β_{X2}	0.2263	0.0512	0.3989	0.0889	0.2260	0.0502	0.3988	0.0872
β_{X3}	0.3203	0.0432	0.3864	0.0713	0.3207	0.0436	0.3884	0.0705
β_{X4}	0.3225	0.0398	0.3912	0.0648	0.3228	0.0404	0.3923	0.0651
β_{X5}	0.2148	0.0422	0.0896	0.0743	0.2149	0.0421	0.0906	0.0722
β_{X6}	0.0930	0.0328	0.1402	0.0538	0.0929	0.0328	0.1401	0.0538
β_{X7}	-0.4785	0.0615	-0.0768	0.1104	-0.4781	0.0595	-0.0780	0.1132
β_{X8}	-0.3897	0.0628	-0.0832	0.1149	-0.3892	0.0609	-0.0834	0.1139
β_{X9}	0.1933	0.0393	0.0943	0.0644	0.1932	0.0397	0.0942	0.0656
β_{X10}	0.2037	0.0609	0.2121	0.0966	0.2031	0.0612	0.2129	0.0959
β_{X11}	-0.2048	0.0519	-0.1349	0.0870	-0.2055	0.0529	-0.1378	0.0897
α_ℓ	0.6853	0.0648	0.6776	0.1728	0.6853	0.0647	0.6766	0.1626
ω	2.0703				2.0489			
LogLik	-25,531.98				-25,532.55			
AIC	51,117.96				51,119.10			

Table 5.4: Parameter estimation - Sarmanov Approximation for the pair COL/COM

	Sarmanov				Approximated Sarmanov			
	AF		NAF		AF		NAF	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-3.3217	0.1662	-3.2043	0.1457	-3.3188	0.1280	-3.2001	0.1191
β_{X1}	0.7385	0.1263	0.5423	0.1118	0.7391	0.1067	0.5414	0.1041
β_{X2}	0.1613	0.0814	0.2754	0.0716	0.1616	0.0718	0.2748	0.0671
β_{X3}	0.3642	0.0642	0.2962	0.0566	0.3669	0.0647	0.2992	0.0563
β_{X4}	0.3927	0.0590	0.2724	0.0524	0.3941	0.0592	0.2738	0.0523
β_{X5}	0.3032	0.0617	0.1504	0.0559	0.3049	0.0624	0.1521	0.0558
β_{X6}	0.1252	0.0487	0.0662	0.0428	0.1249	0.0483	0.0660	0.0426
β_{X7}	-0.5709	0.0903	-0.4073	0.0897	-0.5728	0.0843	-0.4088	0.0786
β_{X8}	-0.4594	0.0932	-0.3316	0.0887	-0.4596	0.0868	-0.3319	0.0810
β_{X9}	0.2146	0.0578	0.1742	0.0514	0.2148	0.0575	0.1737	0.0520
β_{X10}	0.1528	0.0925	0.2425	0.0770	0.1551	0.0908	0.2431	0.0769
β_{X11}	-0.2175	0.0917	-0.2102	0.0722	-0.2220	0.0755	-0.2154	0.0672
α_ℓ	0.8176	0.3940	0.7122	0.3092	0.7696	0.1301	0.6539	0.1111
ω	4.4886				5.6594			
LogLik	-20,878.76				-20,884.45			
AIC	41,811.52				41,822.90			

Table 5.5: Parameter estimation - Sarmanov Approximation for the pair AF/NAF

different, at $-20,884.45$. To summarize, the Approximated Sarmanov is not similar to the Sarmanov model, but the approximation seems to be reasonable.

Tables 5.6 and 5.7 show the estimated parameters for the dynamic models presented earlier in the paper: H-F, Bivariate H-F and Dynamic Sarmanov. We observe that the estimated parameters (the intercept and the eleven covariates from Table 5.1) are approximately the same for all (stationary and dynamic) models, which is a condition that shows consistency between models (see for example Gourieroux et al. (1984)).

5.2.1 Specification Tests

All the models presented in this paper that generalize the MVNB model are somewhat related given certain linked parameter restrictions. We illustrated the situation in Figure 5.2.1, where links between nested models are shown. Note that model $DS1$ refers to an intermediary model in the scheme.

This illustration is used to test all the linked models via a likelihood ratio test to check which model to retain. We perform likelihood ratio tests between all nested models used

	Harvey-Fernandes				Bivariate Harvey-Fernandes				Dynamic Sarmanov			
	COL		COM		COL		COM		COL		COM	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-2.5749	0.0921	-4.1856	0.1701	-2.5749	0.0915	-4.2021	0.1687	-2.5739	0.1009	-4.1879	0.1771
β_{X1}	0.6466	0.0773	0.4664	0.1400	0.6465	0.0777	0.4706	0.1486	0.6482	0.0811	0.4681	0.1413
β_{X2}	0.2286	0.0514	0.4007	0.0868	0.2286	0.0509	0.4028	0.0901	0.2294	0.0577	0.4006	0.0878
β_{X3}	0.3219	0.0433	0.3854	0.0721	0.3218	0.0433	0.3824	0.0714	0.3209	0.0434	0.3792	0.0711
β_{X4}	0.3236	0.0400	0.3895	0.0658	0.3236	0.0401	0.3860	0.0662	0.3226	0.0407	0.3864	0.0657
β_{X5}	0.2179	0.0423	0.0881	0.0751	0.2179	0.0424	0.0854	0.0745	0.2174	0.0426	0.0866	0.0758
β_{X6}	0.0907	0.0330	0.1400	0.0543	0.0907	0.0329	0.1453	0.0543	0.0911	0.0333	0.1398	0.0551
β_{X7}	-0.4775	0.0621	-0.0811	0.1138	-0.4775	0.0616	-0.0725	0.1115	-0.4798	0.0626	-0.0789	0.1255
β_{X8}	-0.3892	0.0637	-0.0854	0.1197	-0.3892	0.0622	-0.0799	0.1094	-0.3908	0.0640	-0.0847	0.1268
β_{X9}	0.1932	0.0397	0.0947	0.0648	0.1932	0.0401	0.0918	0.0666	0.1941	0.0398	0.0947	0.0679
β_{X10}	0.2022	0.0613	0.2122	0.0960	0.2022	0.0608	0.2094	0.0964	0.2034	0.0610	0.2115	0.0963
β_{X11}	-0.2067	0.0524	-0.1392	0.0905	-0.2067	0.0521	-0.1396	0.0887	-0.2058	0.0521	-0.1336	0.0935
$\alpha_{\ell,0}$	0.5021	0.0689	0.4472	0.1807	0.5021	0.0692	0.4273	0.1776	0.5027	0.0709	0.4482	0.1781
ν_ℓ	0.7057	0.0539	0.6421	0.1330	0.7057	0.0546	0.6083	0.1155	0.7062	0.0549	0.6240	0.1549
γ_ℓ					0	0.1438	0.2996	0.1611	0	0.2106	0.4162	0.2771
ω									0.7417			
LogLik	-25,521.67				-25,519.06				-25,517.31			
AIC	51,099.34				51,098.12				51,096.62			

Table 5.6: Parameter estimation - Dynamic Models for the pair COL/COM

	Harvey-Fernandes				Bivariate Harvey-Fernandes				Dynamic Sarmanov			
	AF		NAF		AF		NAF		AF		NAF	
	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error	Estimates	std error
β_0	-3.3129	0.1281	-3.1975	0.1184	-3,312953	0.1352	-3.2068	0.1237	-3.3223	0.1289	-3.2097	0.1200
β_{X1}	0.7385	0.1066	0.5413	0.1032	0.7380	0.1067	0.5408	0.1049	0.7420	0.1080	0.5470	0.1046
β_{X2}	0.1618	0.0711	0.2747	0.0651	0.1615	0.0730	0.2779	0.0685	0.1642	0.0743	0.2782	0.0680
β_{X3}	0.3697	0.0638	0.3037	0.0556	0.3660	0.0645	0.2983	0.0564	0.3633	0.0647	0.2962	0.0565
β_{X4}	0.3953	0.0587	0.2760	0.0520	0.3925	0.0593	0.2747	0.0526	0.3925	0.0591	0.2726	0.0525
β_{X5}	0.3072	0.0615	0.1550	0.0554	0.3060	0.0627	0.1514	0.0559	0.3041	0.0627	0.1527	0.0562
β_{X6}	0.1236	0.0478	0.0651	0.0428	0.1238	0.0477	0.0639	0.0428	0.1234	0.0479	0.0647	0.0428
β_{X7}	-0.5739	0.0854	-0.4104	0.0782	-0.5714	0.0892	-0.4072	0.0818	-0.5705	0.0870	-0.4050	0.0818
β_{X8}	-0.4592	0.0883	-0.3318	0.0806	-0.4582	0.0915	-0.3314	0.0828	-0.4595	0.0887	-0.3314	0.0830
β_{X9}	0.2128	0.0579	0.1742	0.0505	0.2115	0.0573	0.1762	0.0513	0.2136	0.0580	0.1754	0.0519
β_{X10}	0.1554	0.0903	0.2451	0.0753	0.1545	0.0905	0.2438	0.0767	0.1508	0.0913	0.2407	0.0771
β_{X11}	-0.2300	0.0738	-0.2207	0.0670	-0.2287	0.0783	-0.2077	0.0701	-0.2182	0.0756	-0.2074	0.0675
$\alpha_{\ell,0}$	0.3975	0.1413	0.2901	0.0998	0.3725	0.1472	0.2962	0.1826	0.5453	0.1589	0.4755	0.1598
ν_ℓ	0.7827	0.1600	0.8090	0.1616	0.7848	0.1613	0.8329	0.1794	0.6939	0.0832	0.6742	0.1264
γ_ℓ					0.7697	0.4015	0.9659	1.1541	0.4296	0.2008	0.4864	0.2118
ω									9.6347			
LogLik	-20,898.77				-20,890.34				-20,870.36			
AIC	41,853.54				41,840.68				41,802.72			

Table 5.7: Parameter estimation - Dynamic Models for the pair AF/NAF

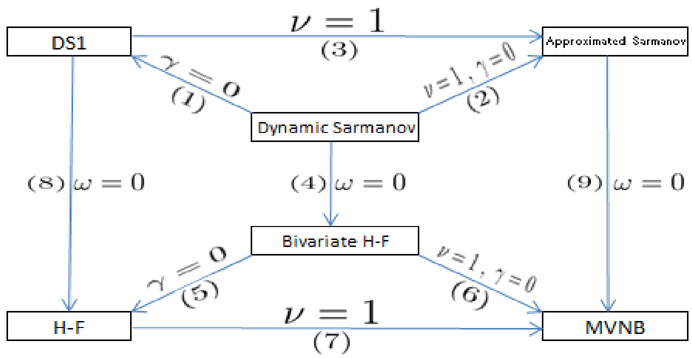

Figure 5.1: Links between the models

in the empirical illustration. As mentioned, corrections to the likelihood ratio test are used when the null hypothesis is on the boundary of the parameter space. Because we have to test many parameters simultaneously, we use the result of Self and Liang (1987), namely that the distribution of the likelihood ratio statistic under the null hypothesis is:

$$F(x) = \sum_{d_f=0}^{k} p_{d_f} F(x; d_f),$$

where $F(x; d_f)$ is the cumulative distribution function of a Chi-Square distribution with d_f degrees of freedom, and p_{d_f} represents the probability of success of a binomial distribution with parameters $n = k$ and $p = 0.5$. The parameter k is the difference of parameters between the null and the alternative hypothesis. For example, for a likelihood ratio test where 4 parameters are tested simultaneously on the boundary of their parameter space, the null distribution can be defined as:

$$\frac{1}{16}0 + \frac{4}{16}\chi^2(1) + \frac{6}{16}\chi^2(2) + \frac{4}{16}\chi^2(3) + \frac{1}{16}\chi^2(4).$$

We summarize the main results of these tests in Table 5.8.

For the pair COM/COL, we can observe that all forms of dependence between comprehensive and collision claims are rejected. Indeed, compared with the independence case, all alternative models supposing dependence between the types of claims are rejected. However, following Table 5.7, we see that in a dynamic setting, collision claims might provide insight into comprehensive claims prediction. In fact, we observe that γ_2 is significant, meaning

[1]Corrected likelihood ratio test: the null hypothesis is on the boundary of the parameter space.

Test	Alternative model	Null model	DF	COL/COM Test statistic	COL/COM p-value	AF/NAF Test statistic	AF/NAF p-value
(2)[1]	Dynamic Sarma.	Approx. Sarma.	4	30.48	< 0.001	28.18	< 0.001
(4)	Dynamic Sarma.	Biv. H-F	1	3.50	0.061	39.96	< 0.001
(5)[1]	Biv. H-F	H-F	2	5.22	0.003	16.86	< 0.001
(6)[1]	Biv. H-F	MVNB	4	27.46	< 0.001	19.00	< 0.001
(7)[1]	H-F	MVNB	2	22.24	< 0.001	2.14	0.097
(9)	Approx. Sarma.	MVNB	1	0.48	0.488	30.78	< 0.001

Table 5.8: Specification Tests

that collision loss experience could enhance comprehensive claim prediction. We also see that $\hat{\gamma}_1 = 0$, meaning that no information is brought from the comprehensive claims to collision claims prediction. This is also confirmed by test (5), where a Bivariate H-F is not rejected against an H-F model, meaning that borrowing information from the correlated type of claim could improve the prediction. However, tests indicate that all stationary models are rejected against dynamic models. This shows that the data favor a model allowing greater weight to the most recent claims. Finally, when we compare the p-values between tests (4) and (9), we observe that the dependence parameter ω is becoming much more significant in a dynamic context.

For the pair AF/NAF, we observe, overall, very significant dependence between AF and NAF claims. When comparing stationary and dynamic models, we see that the tests reject stationary models in favor of dynamic models, meanwhile the MVNB is not rejected over the H-F model. Interestingly, this means that for the pair AF/NAF, a dynamic model is preferred when the information of the correlated type of claim is incorporated. Thus, the prediction is improved when additional information of the other type of claims is added to the model, which justifies and supports the intuition of adding the parameter γ_ℓ. Finally, note that by introducing a dependence parameter ω in the model, the values of $\hat{\gamma}_\ell$ changed considerably. Indeed, while $\hat{\gamma}_{AF}$ was equal to 0.7697 for the Bivariate H-F model, it goes down to 0.4296 for the Dynamic Sarmanov model. This means that NAF loss experience has a greater impact on the AF premium with a Bivariate H-F model than with a Dynamic Sarmanov model. We think that this can be explained by the flexibility induced by the ω parameter in the Dynamic Sarmanov model, where this extra parameter can be used to model the variance independently from the mean.

5.3 Premium Comparison

Each of the models presented in this paper has different properties, and generates different *a priori* and predictive premiums. Beside comparisons of the fit of the model to empirical data, it is useful to compare the premiums. For illustration purposes, we consider three different profiles classified as good, average and bad drivers, given their risk characteristics. The selected profiles are described in Table 5.9 and their respective *a priori* premiums are

Profile Number	Type of Profile	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}
1	Good	0	1	1	0	0	0	1	0	0	0	1
2	Average	0	1	0	1	0	0	0	1	0	0	0
3	Bad	1	0	0	0	0	0	0	1	0	0	0

Table 5.9: Profiles Analyzed

	Good Profile			Average Profile			Bad Profile		
	Mean		Variance	Mean		Variance	Mean		Variance
Models	COL	COM	Total	COL	COM	Total	COL	COM	Total
MVNB	0.0665	0.0268	0.1008	0.0894	0.0307	0.1332	0.0981	0.0222	0.1351
Sarmanov	0.0665	0.0268	0.1015	0.0894	0.0306	0.1342	0.0981	0.0221	0.1358
Approx. Sarma	0.0665	0.0268	0.1015	0.0895	0.0307	0.1344	0.0981	0.0222	0.1359
H-F	0.0666	0.0268	0.1039	0.0896	0.0308	0.1386	0.0985	0.0223	0.1412
Bivariate H-F	0.0666	0.0265	0.1037	0.0896	0.0304	0.1382	0.0985	0.0221	0.1411
Dynamic Sarma	0.0666	0.0268	0.1042	0.0896	0.0306	0.1388	0.0986	0.0223	0.1418

Table 5.10: *A priori* premiums for the pair COL/COM

given in Tables 5.10 and 5.11. These tables show that the values exhibit small differences for the six most useful models presented in this paper. We observe the same trend for variance, with a slight increase for the dynamic models compared with the stationary ones. These results are not surprising because all models have the same form of expected values and, as specified in the previous section, all estimates of βs are similar.

We expect more differences for the predictive premiums, because some models are dynamic, others depend only on past claims and still other models also depend on the claim experience of the other type of claims. For illustration purposes, we use the pair AF/NAF only. We have kept the estimated parameters of the *a priori* analysis and projected a loss experience of 10 years for a medium-risk profile. Although other situations can easily be illustrated, because closed-form formulas have been found to compute the predictive premiums for each model, we focus here on five specific situations. Table 5.12 provides a detailed

	Good Profile			Average Profile			Bad Profile		
	Mean		Variance	Mean		Variance	Mean		Variance
Models	AF	NAF	Total	AF	NAF	Total	AF	NAF	Total
MVNB	0.0277	0.0387	0.0723	0.0401	0.0509	0.1017	0.0480	0.0504	0.1104
Sarmanov	0.0277	0.0388	0.0704	0.0397	0.0504	0.0972	0.0477	0.0501	0.1060
Approx. Sarma	0.0277	0.0388	0.0709	0.0398	0.0506	0.0986	0.0479	0.0503	0.1077
H-F	0.0277	0.0388	0.0736	0.0402	0.0509	0.1040	0.0481	0.0504	0.1131
Bivariate H-F	0.0277	0.0389	0.0739	0.0401	0.0505	0.1035	0.0482	0.0499	0.1127
Dynamic Sarma	0.0278	0.0389	0.0743	0.0398	0.0503	0.1039	0.0478	0.0501	0.1140

Table 5.11: *A priori* premiums for the pair AF/NAF

	Exp. #1		Exp. #2		Exp. #3		Exp. #4		Exp. #5	
Year	AF	NAF	AF	NAF	AF	NAF	AF	NAF	AF	NAF
1	0	0	1	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	1	0	0
3	0	0	1	1	0	0	0	1	0	0
4	0	0	0	0	0	0	0	0	0	0
5	0	0	1	1	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	1
7	0	0	0	0	1	1	0	0	0	1
8	0	0	0	0	1	0	0	0	0	0
9	0	0	0	0	0	1	0	0	0	1
10	0	0	0	0	1	0	0	0	0	0

Table 5.12: Various 10-Year Loss Experiences

description of these loss experience situations. The first loss experience describes a claim-free situation. The second loss experience illustrates the situation of an insured with old claims for both coverages. The third experience is a situation with recent claims for both types of claims. Finally, the fourth and fifth loss experiences correspond to claim-free situation for AF coverage, while old claims (Exp. #4) or recent claims (Exp. #5) are considered for NAF coverage.

The computed predictive premiums are presented in Tables 5.13 and 5.14 . For the claim-free situation of loss experience #1, the predicted premiums of the dynamic models are much lower than for the stationary models. We also observe the same trend for the situation where insured claimed three times in the first five years, but showed a neat progression in the most recent years. This is expected given that dynamic models have an extra parameter ν that allows us to weight past claims. For example, to compute next year's premium using a dynamic model with ν approximately equal to 70%, we would assign a claim that happens in the previous year a weight of 100%, a claim 5 years old a weight of 24%, and a claim 10 years old a weight of only 4% on the predictive premium. Meanwhile, for static models, each claim weights 100% in the calculation of the premium regardless of the occurrence time. This highlights an interesting feature of the dynamic models, where an insurer using a dynamic model in its ratemaking system would reward the positive evolution of its insured's claim experience. The insured would therefore be encouraged to improve his profile in the future even if he had more claims in the past. On the other hand, the dynamic models compensate the low premiums of those first two situations by offering higher premiums for insureds with recent claims. We also observe the same trend for predictive variance, with a more significant difference between stationary and dynamic models.

The dependence between claim types can be studied in a similar way by analyzing predictive premiums of the two types of claims simultaneously. We know that models that suppose independence between claim types do not considerer the claim experience of the other type

	Exp. #1			Exp. #2			Exp. #3		
	Mean		Variance	Mean		Variance	Mean		Variance
Models	AF	NAF	Total	AF	NAF	Total	AF	NAF	Total
MVNB	0.0221	0.0206	0.0450	0.1565	0.1988	0.3742	0.1565	0.1988	0.3742
Sarma	0.0230	0.0261	0.0513	0.1308	0.1616	0.3038	0.1308	0.1616	0.3038
Approx. Sarma.	0.0262	0.0285	0.0571	0.1284	0.1594	0.2970	0.1284	0.1594	0.2970
H-F	0.0067	0.0065	0.0159	0.1222	0.1496	0.3243	0.4203	0.3404	0.9082
Bivariate H-F	0.0039	0.0051	0.0102	0.1213	0.1808	0.3366	0.3776	0.4105	0.8785
Dynamic Sarma.	0.0029	0.0023	0.0063	0.0795	0.0831	0.2173	0.4656	0.4463	1.0825

Table 5.13: Predictive Premiums for the pair AF/NAF (1)

of claims in the computation of the predictive premiums. The AF premiums calculated by the MVNB and the H-F models illustrate this situation. Indeed, the AF premium does not depend on the NAF loss experience, because the AF premium is the same for loss experiences #1, #4 and #5. In contrast, the AF premium of the Sarmanov model, which allows for dependence between claim types, shows that the loss experience of the NAF coverage has an impact. Indeed, the AF premium is different between loss experiences #1 and #4. However, the premium is the same for loss experiences #4 and #5, because the Sarmanov model is static. It is interesting to see that the AF premiums of the Approximated Sarmanov model do not behave the same way as in the Sarmanov model. Indeed, we cannot observe differences between AF premiums for loss experiences #1, #4 and #5. As explained in Section 4.2.2, this comes from the construction of the Approximated Sarmanov model.

Finally, it is interesting to analyze the premiums of the Bivariate H-F and the Dynamic Sarmanov models. Both models allow past claims experience of NAF coverage to affect the AF premium. We see clear differences between the premiums of loss experiences #1, #4 and #5. Another striking observation is the difference between the computed premiums of each model. This can be explained straightforwardly by looking at the estimated parameters $\hat{\gamma}_{AF}$ of Table 5.7, which we analyzed earlier. Lastly, we observe a difference in variances between the Dynamic Sarmanov and Bivariate Harvey-Fernandes model, due to the addition of the dependence parameter ω as discussed above.

6 Concluding Remarks

Panel data models for claims count are used to model the potential dependence between the number of claims of contracts of the same insured. A generalization into bivariate panel data models can illustrate dependence between coverages. A dynamic approach allows the most recent claims to be more predictive than oldest ones in the prediction. In this paper, we proposed a new model that captures all these features of the panel data models for claims count.

	Exp. #4			Exp. #5		
	Mean		Variance	Mean		Variance
Models	AF	NAF	Total	AF	NAF	Total
MVNB	0.0221	0.1988	0.2338	0.0221	0.1988	0.2338
Sarma	0.0439	0.1501	0.2026	0.0439	0.1501	0.2026
Approx. Sarma.	0.0262	0.1594	0.1914	0.0262	0.1594	0.1914
H-F	0.0067	0.1496	0.1860	0.0067	0.3404	0.4128
Bivariate H-F	0.0495	0.1044	0.1714	0.1241	0.2114	0.3736
Dynamic Sarma.	0.0213	0.0640	0.1097	0.0822	0.3010	0.4672

Table 5.14: Predictive Premiums for the pair AF/NAF (2)

The Sarmanov family of multivariate distribution has been used to model the joint density of the random effects. We show that the form of the posterior density of this family of distributions is almost the same form as that of the prior density. To be able to use the dynamic approach proposed by Bolancé et al. (2007), an approximation of the posterior density has been made. We showed that the approximation is reasonnable, but not identical to the Sarmanov model. The Approximated Sarmanov model allowed us to construct a Dynamic Sarmanov model that possesses nice properties: closed-form expressions of the predictive distribution and closed-form expressions of the predictive premium.

We implemented the model with a sample of insurance data that comes from a major Canadian insurance company. The empirical illustration has been performed on two pairs of claim types: collision vs comprehensive and at-fault vs non-at-fault collision claims, which allows us to expose the proposed model to a wider range of situations. For each pair of coverage, a dynamic structure seemed to be relevant, the Dynamic Sarmanov model was one of the best models to adjust the data.

The Dynamic Sarmanov has been applied to a Poisson-gamma structure, but other combinations are easily possible (Poisson-Inverse Gaussian, NB2-Beta, etc.), as long as a conjugate property can be found. Also, the proposed approach can easily be generalized to more than two lines of business. Indeed, it is possible to extend the Sarmanov family of distributions to the multivariate case. Based on our data, a triplet of claim types using comprehensive, at-fault and non-at-fault claims could be interesting for future research. The trivariate Sarmanov joint density would be expressed as:

$$\begin{aligned} u^S(\theta_1, \theta_2, \theta_3) &= h(\theta_1, \alpha_1, \tau_1)\, h(\theta_2, \alpha_2, \tau_2)\, h(\theta_3, \alpha_3, \tau_3) \\ &\quad \times \left(1 + \omega_{12}\phi_1\phi_2 + \omega_{13}\phi_1\phi_3 + \omega_{23}\phi_2\phi_3 + \omega_{123}\phi_1\phi_2\phi_3\right). \end{aligned}$$

The correlation structure is expensive in terms of parameters: the model supposes four parameters to model dependence. It would be interesting to understand how each parameter affects dependence between claim types. Moreover, a multivariate model could also be performed to incorporate claim severity analysis.

References

Abdallah, A., Boucher, J.-P., Cossette, H., and Trufin, J. (2015). Sarmanov Family of Bivariate Distributions for Multivariate Loss Reserving Analysis. *North American Actuarial Journal. To be published.*

Bolancé, C., Bahraoui, Z., and Artís, M. (2014). Quantifying the Risk Using Copulae with Nonparametric Marginal. *Insurance: Mathematics and Economics*, 58:46–56.

Bolancé, C., Denuit, M., Guillén, M., and Lambert, P. (2007). Greatest Accuracy Credibility with Dynamic Heterogeneity: the Harvey-Fernandes model. *Belgian Actuarial Bulletin*, 1:14–18.

Boucher, J.-P., Denuit, M., and Guillen, M. (2007). Risk Classification for Claim Counts: A Comparative Analysis of Various Zero-Inflated Mixed Poisson and Hurdle Models. *North American Actuarial Journal*, 11(4):110–131.

Boucher, J.-P., Denuit, M., and Guillen, M. (2008). Models of Insurance Claim Counts with Time Dependence Based on Generalisation of Poisson and Negative Binomial Distributions. *Variance*, 2(1):135–162.

Boucher, J.-P. and Inoussa, R. (2014). A Posteriori Ratemaking with Panel Data. *ASTIN Bulletin*, 44(3):587–612.

Cohen, L. (1984). Probability Distributions With Given Multivariate Marginals. *Journal of Mathematical Physics*, 25:2402–2403.

Danaher, P. and Smith, M. (2011). Modeling Multivariate Distributions Using Copulas: Applications in Marketing. *Marketing Science*, 30 (1):4–21.

Denuit, M., Marechal, X., Pitrebois, S., and Walhin, J.-F. (2007). Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Scales. *Wiley: New York*.

Frees, E. and Valdez, E. (2008). Hierarchical Insurance Claims Modeling. *Journal of the American Statistical Association*, 103 (484):1457–1469.

Gerber, H. and Jones, D. (1975). Cedibility Formulas of the Updating Type. *In "Credibility", edited by Kahn, P.M.*, Academic Press, New York:89–105.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo-Maximum Likelihood Methods: Theory. *Econometrica*, 52(4):681—700.

Harvey, A. and Fernandes, C. (1989). Time Series Models for Count or Qualitative Observations. *Journal of Business and Economics Statistics*, 7:407–422.

Hernández-Bastida, A. and Fernández-Sánchez (2012). A Sarmanov Family with Beta and Gamma Marginal Distributions: An Application to the Bayes Premium in a Collective Risk Model. *Statistical Methods and Applications*, 21 (4):91–409.

Jewell, W. (1975). The Use of Collateral Data in Credibility Theory: A Hierarchical Model. *Giornale dell Istituto Italiano degli Attuari*, 38:1–16.

Jung, R. and Liesenfeld, R. (2001). Estimating Time Series Models for Count Data Using Effcient Importance Sampling. *Allgemeines Statistisches Archiv*, 85:387–407.

Lee, M.-L. T. (1996). Properties and Applications of the Sarmanov Family of Bivariate Distributions. *Comm. Statist. Theory Methods*, 25(6):1207–1222.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. ISBN 0-4123176-0-5. Chapman and Hall, Boca Raton, FL, 2nd edition.

Miravete, E. (2009). Multivariate Sarmanov Count Data Models. *CEPR DP*, 7463.

Pinquet, J. (1998). Designing Optimal Bonus-Malus Systems From Different Types of Claims. *ASTIN Bulletin*, 28:205–229.

Pinquet, J., Guillen, M., and Bolance, C. (2001). Allowance for Age of Claims in Bonus-Malus Systems. *ASTIN Bulletin*, 31:337–348.

Purcaru, O. and Denuit, M. (2002). On the Dependence Induced by Frequency Credibility Models. *Belgian Actuarial Bulletin*, 2(1):73–79.

Sarmanov, O. (1966). Generalized Normal Correlation and Two-dimensional Frechet Classes. *Doklady (Soviet Mathematics)*, Tom 168:596–599.

Schweidel, D., Fader, P., and Bradlow, E. (2008). A Bivariate Timing Model of Customer Acquisition and Retention. *Marketing Science*, 27 (5):829–843.

Self, S. G. and Liang, K. Y. (1987). Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions. *Journal of the American Statistical Association*, 82(398):605–610.

Shi, P. and Valdez, E. (2014). Longitudinal Modeling of Insurance Claim Counts Using Jitters. *Scandinavian Actuarial Journal*, 2:159–179.

Xu, S., Jones, R., and Grunwald, G. (2007). Analysis of Longitudinal Count Data with Serial Correlation. *Biometrika Journal*, 49(3):416–428.