Hierarchical random effects model for insurance pricing of vehicles belonging to a fleet

Denise Desjardins[†], Georges Dionne[‡], Yang Lu[‡]

18 February 2021

Abstract

We propose a hierarchical random effect model for the posterior insurance ratemaking of vehicles belonging to a fleet by allowing random effects for fleet, vehicle, and time. The model is an alternative to the gamma-Dirichlet model of Angers et al (2018), which does not allow for a closed form posterior ratemaking formula. Our theoretical extension derives a simple and tractable closed form ratemaking formula based on a hierarchical random effects specification. We estimate the corresponding econometric model and compute insurance premiums in relation to the past experience of both the vehicle and the fleet. Our econometric model can also be applied to any other dynamic count modeling application with random individual, time, and common effects, such as labor contracting, chirurgical accidents, or any other random event implying principals and many agents.

Keywords: Hierarchical model, vehicle insurance pricing, posterior ratemaking, random effect, hierarchical random effects, panel data.

JEL codes: C23, C25, C55, G22.

Acknowledgments
Yang Lu acknowledges financial support from CNRS, ANR (through Labex MME-DII) as well as Concordia University (through a Start-up grant). Financial support from the SSHRC of Canada and the SAAQ is also acknowledged as well as the contribution of Claire Boisvert.

[†] Denise Desjardins, HEC Montréal. denise.desjardins@cirrelt.ca.
[‡] Corresponding author: Georges Dionne, Canada Research Chair in Risk Management, HEC Montréal. 3000 Chemin Côte-Ste-Catherine, Montreal (Qc) Canada. Tel: 514-340-6596. E-mail: georges.dionne@hec.ca.
[‡] Yang Lu, Paris 13 University and Concordia University. yang.lu@concordia.ca.

1. Introduction

Several researchers have proposed different models to account for correlations that result from panel data. The use of panel-type individual data has become popular in economic, finance, and actuarial science applications since the 1980s. Early applications include the work of Hausman and Wise (1979), Cameron and Trivedi (1986), Hsiao (1986), Baltagi (1995), and Dionne et al (1997). Two ground-breaking contributions with count data applications are the articles by Hausman et al (1984) and Gouriéroux et al (1984). The former proposed a Maximum Likelihood Estimation (MLE) method for obtaining the estimated parameters while the later developed a pseudo-MLE method.

Following these pathbreaking contributions, accident distribution estimations or insurance pricing applications with parametric models became popular. To name a few, we have the contributions of Dionne and Vanasse (1992), Frangos and Vrontos (2001), Purcaru and Denuit (2003), Boucher and Denuit (2006), Boucher et al (2008), Frees and Valdez (2011), and Cameron and Trivedi (2013). Alternatively, Desjardins et al (2001), Pinquet (2013), Fardilha et al (2016), and Pinquet (2020) proposed semiparametric models for insurance applications. Another class of models is the hierarchical credibility approach with random effects in linear models (Norberg, 1986). Boucher and Guillen (2009) and Pinquet (2013) review the panel count data models applied to insurance pricing. To our knowledge, none of these contributions consider individual and firm effects separately even if some of them analyze accidents of vehicles belonging to a fleet.

The first contribution in the literature that proposes a non-linear econometric model for estimating individual and firm effects with panel data is Angers et al (2018). The matching of longitudinal individual and firm data is very important in environments where the observed outcomes are a function of both parties' observable characteristics (individual and firm) and the unobserved actions and characteristics of both parties (moral hazard and adverse selection). For insurance companies, knowing all sources of accidents involving vehicles belonging to a fleet is essential to developing a fair and incentivized pricing scheme that considers the safety behavior of each actor. This is also important for the regulator, who has to compute the optimal fines for different infractions (driver, fleet owner) that affect accident distributions. Angers et al (2018) extend the parametric model of Hausman et al (1984) to add a firm effect to the individual and time effects in the estimation of event distributions and apply their model to the accident distributions of trucks belonging to fleets of vehicles. They estimate the distribution of vehicle accidents for different fleets over time by first decomposing the explanatory factors into heterogeneous factors linked to vehicles and their drivers, then into heterogeneous factors linked to fleets and their owners, and finally into residual factors.

Factors linked to vehicles and drivers and those linked to fleets and owners can be correlated. For example, a negligent manager may not spend enough money on the mechanical repair of his trucks and might ask his employees to drive fast. However, the employees may also exceed the speed limit without informing the manager. The model of Angers et al (2018) is a three-level hierarchical pricing model with fleet, vehicle, and time effects. Unfortunately, their model does not allow for a closed form posterior ratemaking formula. The insurance price has to be obtained either by Monte-Carlo simulations or approximation formulas (see also Norberg (1986) for credibility approximation formulas).

The goal of this paper is to propose a simple and tractable hierarchical posterior ratemaking model based on hierarchical random effects specifications. Our econometric model can also be applied to any other dynamic count modeling application with random individual and common effects on events involving many agents working for different principals under asymmetric information (Holmstrom, 1982; Laffont and Martimort, 2001).

The rest of the paper is organized as follows. Section 2 introduces the theoretical model. Section 3 computes the theoretical likelihood function. Section 4 derives the posterior insurance pricing formula. Section 5 estimates the econometric model using data on a fleet of vehicles and compares the results to previous contributions in the literature. Section 6 applies the new pricing formula to the data. Section 7 concludes the paper.

2. The theoretical model

Consider I fleets of vehicles, in which each vehicle is doubly indexed by their fleet ID $i = 1,...,I$ and their individual or vehicle ID $j = 1,...,s_i$ within the fleet. Here s_i can be interpreted as the number of vehicles in fleet i, if this number remains constant within fleet i across different periods. In practice, however, this number can change, hence s_i is the total number of vehicles that have belonged to the fleet during any of the T years. In other words, for each given date t, the number of observed vehicles of fleet i is smaller than or equal to s_i. This number may be large, say several dozen or even hundreds. Finally, each vehicle can be observed during several periods T. Thus, the claim counts are triply indexed $X_{i,j,t}$, $i = 1,...,I$, $j = 1,...,s_i$, $t = 1,...,T$, and we denote the associated a prior score with $\lambda_{i,j,t}$, where $\lambda_{i,j,t}$ is the marginal expectation of $X_{i,j,t}$ given all observable covariates. Angers et al (2006) argue for the importance of incorporating a fleet effect, which captures the fleet risk exposure and how the fleet owner manages the risk of its vehicles, as well as a standard individual vehicle effect, which controls the unobservable risk of the driver of each specific vehicle.

Let us remember that a count variable X follows the negative binomial (NB) distribution $NB(\delta, p)$ with parameters $\delta > 0$ and $p \in [0,1]$ if its probability mass function (p.m.f.) is equal to $p(x) = \frac{\Gamma(x+\delta)}{x!\Gamma(\delta)} (1-p)^{\delta} p^x$. We also denote $\gamma(\delta, c)$ the gamma distribution with the shape parameter δ and scale parameter c; $P(\lambda)$ is the Poisson distribution with parameter λ.

We assume that the joint distribution of the observable claim counts $\left(X_{i,j,k}\right)$ is as follows:

- At the highest hierarchical level, the $(N_i)_i$ are *i.i.d.* random effects following $NB(\delta, \beta c)$ distribution, where $\beta c < 1$.
- At the second level, the $\left(Z_{i,j}\right)$ are also random effects that are conditionally *i.i.d.* with $P\left(\beta\eta_{i,j}\right)$ distribution, where the $\left(\eta_{i,j}\right)_j$ are themselves conditionally *i.i.d.* given N_i with

$\gamma(\delta + N_i, c)$ distribution. In other words, the conditional distribution of $Z_{i,j}$ given N_i is $NB\left(\delta + N_i, \frac{\beta c}{1+\beta c}\right)$. Their marginal distributions are $NB(\delta, \beta c)$.

- At the third level, the random effects are $\left(\theta_{i,j,t}\right)_t$ conditionally *i.i.d.* given $Z_{i,j}$, with gamma distribution $\gamma(\delta^* + \beta^* Z_{i,j}, c^*)$.
- Finally, given $\left(\theta_{i,j,t}\right)$, the claim counts $X_{i,j,t}$ are independent and Poisson $P\left(\theta_{i,j,t}\lambda_{i,j,t}\right)$ distributed.

To summarize, Figure 1 presents the model's chain rule:

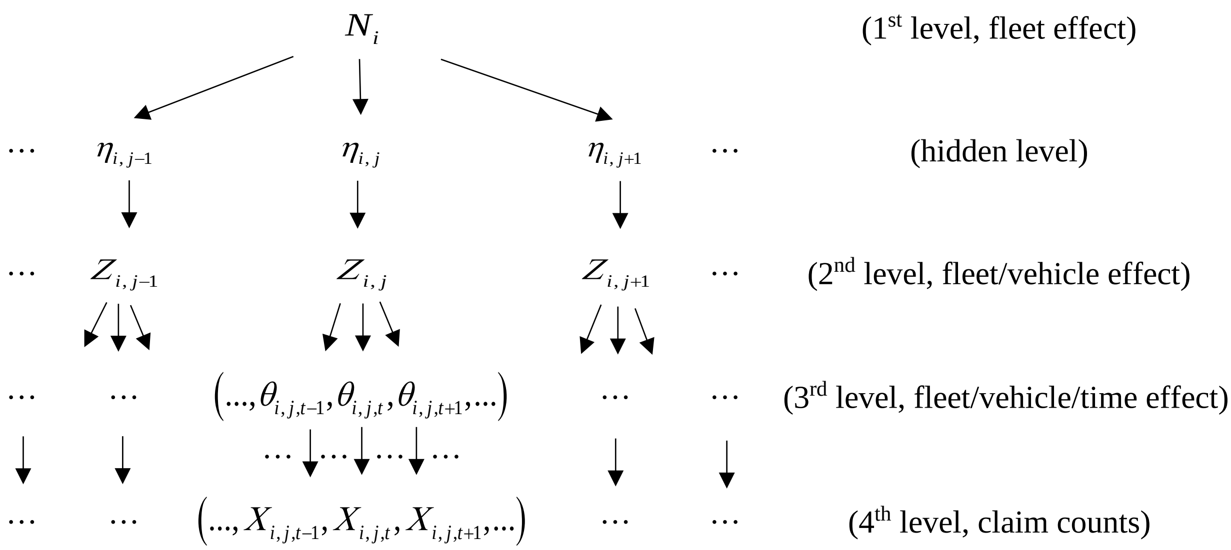

Figure 1: Chain rule of the hierarchical model

We now comment on the state space of these random variables' levels. At the third level, we use $\theta_{i,j,t}$, which are continuously valued, so that $X_{i,j,t}$ have the standard Poisson random effect specification (Dionne and Vanasse, 1989, 1992). The two upper level random effects N_i and $Z_{i,j}$ are both count valued, and it will be shown in Sections 3 and 4 that the discreteness of N_i and $Z_{i,j}$ is essential for the tractability of both the likelihood function and the posterior ratemaking function. Between the two upper levels of count random variables N_i and $Z_{i,j}$, we have introduced a hidden level $\eta_{i,j}$ that is continuously valued. The latter merely serves as an auxiliary mixing variable that allows us to define the conditional distribution of $Z_{i,j}$ given N_i as a Poisson-gamma mixture. Finally, between different levels, we alternate between continuous and count variables by using conditional Poisson and gamma distributions. This technique is well known in the time series literature (Pitt and Walker, 2005; Gouriéroux and Lu, 2019) and has the advantage of leading to relatively tractable marginal and conditional distributions, which are summarized in the following proposition:

Proposition 1

1. *The marginal distribution of* $\eta_{i,j}$ *is* $\gamma\left(\delta,\frac{c}{1-\beta c}\right)$, *whereas the marginal distribution of* $Z_{i,j}$ *is* $NB(\delta,\beta c)$, *and the correlation coefficient between* $Z_{i,j}$ *and* $Z_{i,j+1}$ *is* $Corr\left[Z_{i,j},Z_{i,j+1}\right]=\beta c$.

2. *The marginal distribution of* $\theta_{i,j,k}$ *is generically not gamma, except if* $\delta^*=\delta$ *and* $\beta^*=1$. *In this case* $\theta_{i,j,k}$ *has the* $\gamma\left(\delta,\frac{c^*}{1-\beta c}\right)$ *distribution.*

3. *The correlation coefficient at the lowest level is:*

$$Corr\left[\eta_{i,j,t_{-1}},\eta_{i,j,t}\right]=\frac{(\beta^*)^2\frac{\delta\beta c}{(1-\beta c)^2}}{\beta^*\frac{\delta\beta c}{(1-\beta c)^2}+\delta^*+(\beta^*)^2\frac{\delta\beta c}{1-\beta c}}. \tag{1}$$

Proof. Properties 1 and 2 are a direct consequence of the Poisson-gamma conjugacy, reviewed in Appendix 1. As for Property 3, we have, from the (co)variance decomposition formula:

$$\begin{aligned}V\left[\theta_{i,j,t}\right]&= E\left[V\left[\theta_{i,j,t}\mid Z_{i,j}\right]\right]+V\left[E\left[\theta_{i,j,t}\mid Z_{i,j}\right]\right]\\&=(c^*)^2\left[\delta^*+\beta^*\delta\ \frac{\beta c}{1-\beta c}\right]+(c^*)^2(\beta^*)^2\ \frac{\delta\beta c}{1-\beta c}\\Cov\left[\theta_{i,j,t},\theta_{i,j,t+1}\right]&=(c^*)^2(\beta^*)^2\ \frac{\delta\beta c}{1-\beta c}\ .\end{aligned}$$

□

As expected, the correlation coefficient in (1) does not depend on the scale parameter c^* and, in practice, the scale parameter c^* at the lowest level can be chosen such that $E\left[\theta_{i,j,t}\right]=1$, that is:

$$\left[\delta^* + \delta\beta^* \frac{\beta c}{1-\beta c}\right] c^* = 1. \tag{2}$$

We can distinguish between different values of β^* and δ^*. Three special cases are worth mentioning:

- When $\delta^* = \delta$ and $\beta^* = 1$, we recover the value βc for the correlation coefficient in (1). This implies, in this special case, that the correlation at the lowest level, $Corr\left[\theta_{i,j,t-1}, \theta_{i,j,t}\right]$, and that, at the intermediate level, $Corr\left[\eta_{i,j,t-1}, \eta_{i,j,t}\right]$, are both equal to βc. This case might be too restrictive to be applied, however;
- When β^* goes to infinity and δ^* remains fixed, the correlation attains its maximum at 1;
- When δ^* goes to infinity and β^* remains fixed, or when β^* goes to zero and δ^* remains fixed, the correlation goes to zero, which is its minimum value.

Thus, by allowing for arbitrary positive values for β^* and δ^*, the correlation coefficient $Corr\left[\eta_{i,j,t-1}, \eta_{i,j,t}\right]$ can attain any values in [0,1]. In the limiting case where the correlation attains 1, we get a model with time-invariant random effects $\theta_{i,j,t}$; in the other limiting case where the correlation attains 0, we get a model with independent random effects, that is, with neither fleet effect nor individual effect.

3. The likelihood function

In this section we compute the likelihood function of the model. We can write:

$$\ell\left(\left(X_{i,j,t}\right)_{i,j,t}\right) = \prod_{i=1}^{I} E\left[\prod_{j=1}^{s_i}\prod_{t=1}^{T} P\left[X_{i,j,t} = x_{i,j,t}, \forall i,j,t \,|\, \left(\theta_{i,j,t}\right), \left(Z_{i,j}\right), N_i\right]\right] \tag{3}$$

$$= \prod_{i=1}^{I} E\left[E\left[\prod_{j=1}^{s_i} \prod_{t=1}^{T} \frac{e^{-\lambda_{i,j,t}\theta_{i,j,t}} \left(\lambda_{i,j,t}\theta_{i,j,t}\right)^{x_{i,j,t}}}{x_{i,j,t}!} \Big| \left(Z_{i,j}\right), N_i \right]\right] \quad (4)$$

$$= \prod_{i=1}^{I} \left\{ \left[\prod_{j=1}^{s_i} \prod_{t=1}^{T} \frac{\left(\lambda_{i,j,t}\right)^{x_{i,j,t}}}{x_{i,j,t}!} \right] E\left[E\left[\prod_{j=1}^{s_i} \prod_{t=1}^{T} \frac{\Gamma\left(\delta^* + \beta^* Z_{i,j} + x_{i,j,t}\right)}{\Gamma\left(\delta^* + \beta^* Z_{i,j}\right)} \frac{\left(c^*\right)^{x_{i,j,t}}}{\left(1 + c^* \lambda_{i,j,t}\right)^{\delta^* + \beta^* Z_{i,j} + x_{i,j,t}}} \Big| N_i \right]\right] \right\} \quad (5)$$

where, in equation (3), the conditional probability is a Poisson distribution of $X_{i,j,t}$ given $\theta_{i,j,t}$ and, in equation (4), the inner conditional expectation is with respect to the conditional distribution of $\theta_{i,j,t}$ given $Z_{i,j}$. In both equations the outer expectation is with respect to the distribution of all the latent random variables such as $\theta_{i,j,t}$, $Z_{i,j}$ and N_i.

Then we can compute the expectation:

$$M_i\left(X\right) := E\left[E\left[\prod_{j=1}^{s_i} \prod_{t=1}^{T} \frac{\Gamma\left(\delta^* + \beta^* Z_{i,j} + x_{i,j,t}\right)}{\Gamma\left(\delta^* + \beta^* Z_{i,j}\right)} \frac{(c^*)^{x_{i,j,t}}}{\left(1 + c^* \lambda_{i,j,t}\right)^{\delta^* + \beta^* Z_{i,j} + x_{i,j,t}}} \mid N_i \right]\right]$$

where we use the subscript *i* to indicate that this quantity is fleet dependent, and we use the symbol *X* to indicate the fact that $M_i\left(X\right)$ depends on all the observable counts $\left(X_{i,j,t}\right)_{j,t}$, for all vehicles *j* and time *t*.

In this expression, the outer expectation is with respect to the law of N_i, whereas the inner conditional expectation is with respect to the conditional joint distribution of all the $Z_{i,j}$, *j* varying, given N_i. Because these $\left(Z_{i,j}\right)_j$ are conditionally independent given N_i, we can interchange the product operator and the inner conditional expectation, and instead compute:

$$M_i\left(X\right) = E\left[\prod_{j=1}^{s_i} E\left[\prod_{t=1}^{T} \frac{\Gamma\left(\delta^* + \beta^* Z_{i,j} + x_{i,j,t}\right)}{\Gamma\left(\delta^* + \beta^* Z_{i,j}\right)} \frac{(c^*)^{x_{i,j,t}}}{\left(1 + c^* \lambda_{i,j,t}\right)^{\delta^* + \beta^* Z_{i,j} + x_{i,j,t}}} \mid N_i \right]\right]. \quad (6)$$

Then for each $j=1,\ldots,s_i$, the inner expectation in (6) can be expressed as:

$$M_{i,j}(X,N_i):=E\left[\prod_{t=1}^{T}\frac{\Gamma\left(\delta^*+\beta^*Z_{i,j}+x_{i,j,t}\right)}{\Gamma\left(\delta^*+\beta^*Z_{i,j}\right)}\frac{(c^*)^{x_{i,j,t}}}{\left(1+c^*\lambda_{i,j,t}\right)^{\delta^*+\beta^*Z_{i,j}+x_{i,j,t}}}\mid N_i\right]$$
$$=\sum_{z=0}^{\infty}\left[\frac{\Gamma\left(\delta+N_i+z\right)}{\Gamma\left(\delta+N_i\right)z!}\frac{\left(\beta c\right)^z}{\left(1+\beta c\right)^{\delta+N_i+z}}\prod_{t=1}^{T}\frac{\Gamma\left(\delta^*+\beta^*z+x_{i,j,t}\right)}{\Gamma(\delta^*+\beta^*z)}\frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^*z+x_{i,j,t}}}\right]$$

where the summation is with respect to *z*, that is all the possible values of $Z_{i,j}$, and the term $\frac{\Gamma\left(\delta+N_i+z\right)}{\Gamma\left(\delta+N_i\right)z!}\frac{\left(\beta c\right)^z}{\left(1+\beta c\right)^{\delta+N_i+z}}$ is the conditional p.m.f. of $Z_{i,j}$ given N_i, which is $NB\left(\delta+N_i,\frac{\beta c}{1+\beta c}\right)$. Then we can truncate this infinite summation at a sufficiently high order (say, *K*) and get the approximation:

$$M_{i,j}(X,N_i)\approx\sum_{z=0}^{K}\frac{\Gamma\left(\delta+N_i+z\right)}{\Gamma\left(\delta+N_i\right)z!}\frac{\left(\beta c\right)^z}{\left(1+\beta c\right)^{\delta+N_i+z}}\prod_{t=1}^{T}\frac{\Gamma\left(\delta^*+\beta^*z+x_{i,j,t}\right)}{\Gamma(\delta^*+\beta^*z)}\frac{(c^*)^{x_{i,j,t}}}{\left(1+c^*\lambda_{i,j,t}\right)^{\delta^*+\beta^*z+x_{i,j,t}}}. \quad (7)$$

Finally, equation (6) becomes:

$$M_i(X)=E\prod_{j=1}^{s_i}M_{i,j}(X,N_i)$$
$$=\sum_{n=0}^{\infty}\left[\frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!}\left(\beta c\right)^n\left(1-\beta c\right)^{\delta}\prod_{j=1}^{s_i}M_{i,j}(X,n)\right] \quad (8)$$
$$\approx\sum_{n=0}^{K}\left[\frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!}\left(\beta c\right)^n\left(1-\beta c\right)^{\delta}\prod_{j=1}^{s_i}M_{i,j}(X,n)\right]$$

Using the approximation (7) for $\prod_{j=1}^{s_i}M_{i,j}(X,N_i)$ and (8), we get an approximation for $M_i(X)$, which in turn leads to an approximation of the likelihood function given in (5). Thus, the set of parameters of the model is easily obtained by maximizing the log-likelihood function. These parameters include $\beta^*,\delta^*,\delta,\beta c,$ as well as the regression coefficients that enter into the a priori score functions, $\lambda_{i,j,t}$. Indeed, parameter c^* is fixed by the normalization constraint (2), whereas the likelihood function depends on β and c only through their product. Hence only βc is identifiable, and not these two parameters separately.

The choice of order *K* in the infinite summations (7) and (8) is the result of a trade off. On the one hand, the larger *K* is, the better the approximation accuracy; on the other hand, the larger *K* is, the more computational effort the method requires. Fortunately, our framework should involve a very limited computational cost, which allows us to take quite large values of *K* and hence attain high approximation quality. Indeed, the approximation of $M_i(X)$ requires us to consider the first *K* +1 smallest possible values of N_i. For each such value *n*, we need to compute $M_{i,j}(X,n)$ in parallel for different *j*. As a consequence, the computation of the contribution of fleet *i* to the likelihood function requires a multiple of $s_i(K+1)^2$ operations, which is relatively easy even for quite large values of *K* and s_i.

Note that for expository purposes, the above likelihood function has been derived under the assumption that all the s_i vehicles are observed for each of the *T* periods for fleet *i*. If in practice some vehicles are only observed for a subset of $\{1,...,T\}$, it suffices to use the convention $x_{i,j,t}=\lambda_{i,j,t}=0,$ and $0^0=1$ for those triplets (i,j,t).

Let us now compare the above model with the hierarchical model of Angers et al (2018), which assumes that counts $X_{i,j,t}$ are conditionally independent and Poisson distributed $P(\lambda_{i,j,t}\theta_{i,j,t})$, where the random effect $\theta_{i,j,t}$ is further decomposed into:

$$\theta_{i,j,t}=\alpha_{1,i}\alpha_{2,i,j}\alpha_{3,i,j,t}$$

where $\alpha_{1,i}$ follows a gamma distribution, and both $(\alpha_{2,i,j})_j$ and $(\alpha_{3,i,j,t})_t$ follow Dirichlet distributions of dimension s_i and *T*, respectively. Details of the model are presented in the online appendix. The major restrictions of this gamma-Dirichlet approach are: i) it involves Dirichlet distributions of dimension s_i, which becomes cumbersome when the fleet is large; ii) the resulting likelihood function does not have closed form expression, except when the number of vehicles s_i is equal to 1 or 2; iii) the Bayesian updating formula, that is the forecast of counts of one new period $T+1$, possibly for a new vehicle, requires the introduction of new Dirichlet distributions of dimensions $T+1$, and s_i+1, and the corresponding updating formula again does not have a closed form formula; iv) these new Dirichlet specifications are not compatible with the ones used for estimation. For example, the normalization conditions for the time effect is $\sum_{t=1}^{T} \alpha_{i,j,t} = 1$ for estimation but becomes $\sum_{t=1}^{T+1} \alpha_{i,j,t} = 1$ for pricing. Such incompatibility renders the interpretation of random effects $\alpha_{i,j,t}$ rather difficult in a pricing exercise and it might lead to arbitrage opportunities; and v) the correlation between the random effects is not as flexible as in the above model.

Finally, we note that a hierarchical model for fleet insurance has also been proposed by Antonio et al (2010). However, their model differs from ours in at least two respects. First, their model is not applicable to fleets with only one vehicle at a certain point in time as they mention. Second, their model does not allow for a closed form likelihood function, nor for a closed form forecasting formula. In particular, their posterior premium formula (see their Table 7) depends on unobservable random effects for time, vehicle, and fleet levels. These formulas are not directly usable, unless the posterior joint distribution of these random effects is recovered using the Bayes rule. The latter task is highly complicated because this joint distribution is high dimensional and, given past claim experiences, random effects are no longer mutually independent, even if they are assumed to be independent in the prior model. The authors have not documented in detail how this high dimensional posterior distribution is sampled using Markov Chain Monte Carlo (MCMC) techniques. As a result, we have not been able to apply their approach empirically to our dataset.

4. Forecasting formula

Because of the discrete latent random effect representation, the model we propose is also very convenient for posterior ratemaking, which is when counts in period $T+1$ need to be forecasted for insurance pricing. Let's first compute the posterior joint distribution of $\left(Z_{i,j}\right)_j$. First, the prior joint distribution of $\left(Z_{i,j}\right)_j$ has the mixture p.m.f. with mixing the variable N_i:

$$p\left(\left(z_j\right)_j\right)=\sum_{n=0}^{\infty}\frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!}\left(\beta c\right)^n\left(1-\beta c\right)^{\delta}\prod_{j=1}^{s_i}\frac{\Gamma\left(\delta+n+z_j\right)}{\Gamma\left(\delta+n\right)z_j!}\frac{\left(\beta c\right)^{z_j}}{\left(1+\beta c\right)^{\delta+n+z_j}},\quad \forall\left(z_j\right)_j\in N^{s_i}. \quad (9)$$

Next, the conditional joint distribution of all claim counts $\left(X_{i,j,t}\right)$ given N_i and $\left(Z_{i,j}\right)_j$ is proportional to:

$$\prod_{j=1}^{s_i}\prod_{t=1}^{T}\frac{\Gamma(\delta^*+\beta^* z_j+x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)}\frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j+x_{i,j,t}}}, \quad (10)$$

which is the product of the conditional, negative binomial p.m.f. of $X_{i,j,t}$ given $\lambda_{i,j,t}$ and $Z_{i,j}$. Thus, using Bayes' formula, the posterior joint distribution of N_i and $\left(Z_{i,j}\right)_j$ is proportional to the product of equations (9) and (10):

$$p\left(\left(z_j\right)_j \mid \left(x_{i,k,t}\right)_{k,t}\right)$$
$$\propto\sum_{n=0}^{\infty}\frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!}\left(\beta c\right)^n\left(1-\beta c\right)^{\delta}\prod_{j=1"}^{s_i}\left[\frac{\Gamma\left(\delta+n+z_j\right)}{\Gamma\left(\delta+n\right)z_j!}\frac{\left(\beta c\right)^{z_j}}{\left(1+\beta c\right)^{\delta+n+z_j}}\prod_{t=1}^{T}\frac{\Gamma(\delta^*+\beta^* z_j+x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)}\frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j+x_{i,j,t}}}\right]$$

(11)

where for expository purposes we have used the simplified notation $\left(x_{i,k,t}\right)_{k,t}$ to indicate that the conditioning set is all the observed claim counts for all vehicles *k* of the fleet *i* during the first *T* periods. The normalization constant is given by the summation of the right-hand side of (11) with respect to $z_1, z_2, \ldots z_{s_i}$ over all the integrals, that is:

$$\sum_{z_1=0}^{\infty}\cdots\sum_{z_{s_i}=0}^{\infty}\text{RHS of equation (11)}$$

$$=\sum_{n=0}^{\infty}\frac{\Gamma(\delta+n)}{\Gamma(\delta)n!}(\beta c)^n(1-\beta c)^{\delta}\prod_{j=1}^{s_i}\sum_{z_j=0}^{\infty}\left[\frac{\Gamma(\delta+n+z_j)}{\Gamma(\delta+n)z_j!}\frac{(\beta c)^{z_j}}{(1+\beta c)^{\delta+n+z_j}}\prod_{t=1}^{T}\frac{\Gamma(\delta^*+\beta^* z_j+x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)}\frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j+x_{i,j,t}}}\right]$$

$$=\sum_{n=0}^{\infty}\frac{\Gamma(\delta+n)}{\Gamma(\delta)n!}(\beta c)^n(1-\beta c)^{\delta}\prod_{j=1}^{s_i}M_{i,j}(X,n), \tag{12}$$

where we have interchanged the infinite summations over $z_j,\ j=1,\ldots,s_i$ and the product over *j* because of the separability of each term in (11) into functions of each individual z_j, for a given *n*. In particular we can check that given $(X_{i,j,t})_{j,t}$, the second level random effects $(Z_{i,j})_j$ are still conditionally independent given N_i, and both $M_{i,j}(X,n)$ and the term between the brackets in equation (11) can be computed in parallel for a different *j*. As a consequence, we also deduce the marginal posterior distribution of each individual $Z_{i,j}$ given $(X_{i,j,t})_{j,t}$:

$$p\left(z_j \mid (x_{i,k,t})_{k,t}\right)$$

$$=\frac{\sum_{n=0}^{\infty}\frac{\Gamma(\delta+n)}{\Gamma(\delta)n!}(\beta c)^n(1-\beta c)^{\delta}\left[\frac{\Gamma(\delta+n+z_j)}{\Gamma(\delta+n)z_j!}\frac{(\beta c)^{z_j}}{(1+\beta c)^{\delta+n+z_j}}\prod_{t=1}^{T}\frac{\Gamma(\delta^*+\beta^* z_j+x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)}\frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j+x_{i,j,t}}}\right]\prod_{k\neq j}^{s_i}M_{i,k}(X,n)}{\sum_{n=0}^{\infty}\frac{\Gamma(\delta+n)}{\Gamma(\delta)n!}(\beta c)^n(1-\beta c)^{\delta}\prod_{k=1}^{s_i}M_{i,k}(X,n)} \tag{13}$$

Again, the computation of the above conditional distribution requires double infinite summation only. That is an infinite summation over *n* and, for each value of *n*, the computation of $M_j(X,n)$ for different *j*, which itself requires a one-dimensional infinite summation. Let us now consider the posterior expected number of claims at period $T+1$.

$$\begin{aligned} E\left[X_{i,j,T+1} \mid (x_{i,k,t})_{k,t}\right] &= \lambda_{i,j,T+1}\, E\left[\theta_{i,j,T+1} \mid (x_{i,k,t})_{k,t}\right] \\ &= \lambda_{i,j,T+1}\, E\left[c^*(\delta^*+\beta^* Z_{i,j}) \mid (x_{i,k,t})_{k,t}\right] \\ &= \lambda_{i,j,T+1}c^*\delta^* + \lambda_{i,j,T+1}\beta^* c^* E\left[Z_{i,j} \mid (x_{i,k,t})_{k,t}\right] \end{aligned} \tag{14}$$

where the conditional expectation $E\left[Z_{i,j} \mid \left(x_{i,k,t}\right)_{k,t}\right]$ can be obtained from the conditional p.m.f. of $Z_{i,j} \mid \left(x_{i,k,t}\right)_{k,t}$ given by equation (13) through:

$$E\left[Z_{i,j} \middle| \left(x_{i,k,t}\right)_{k,t}\right] = \sum_{z_j=0}^{\infty} z_j p(z_j \left(x_{i,k,t}\right)_{k,t}) \tag{15}$$

which again involves double infinite summations only.

We end this section with two final remarks. First, the above formulas hold true both for a vehicle *j* that has already been observed between *t* = 1 and *t* = *T*, and for a new vehicle that enters into the fleet at date *T* +1. Indeed, in the latter case, it suffices to apply the convention that $\lambda_{i,j,t} = x_{i,j,t} = 0$ for all $t = 1,...,T$ and $0^0 = 1$, while, as previously stated, the gamma-Dirichlet model requires the introduction of new Dirichlet distributions of dimensions *T* +1, and s_i +1. Second, as in the estimation section, all the infinite summations involved will in practice be approximated by finite ones by truncating them at a sufficiently high order *K*.

5. Model estimation with accidents data

We have access to the files of the *Société de l'assurance automobile du Québec* (henceforth referred to as the SAAQ) to create the data base over the period from 1991-1998. The SAAQ is in charge of road safety regulations and is the public insurer for bodily injuries linked to traffic accidents.

Our starting point is the whole population of fleets registered in Québec in July 1997. To be registered the fleets must be the owner of at least one truck that is not used for emergencies. In this study we use fleets with at least two trucks. Data on fleets contain information on violations (with convictions) committed by the fleet between 1989 and 1998 and information identifying the fleet.

We can link vehicles to fleets. From the authorization status, we obtain information describing the vehicle. For each plate number, we have data covering the 1990-1998 period drawn from the files on the mechanical inspection of vehicles and from the record of violations with convictions and demerit points for: speeding, failure to stop at a red light or stop sign, illegal passing, etc., and data on all accidents. These include all the traffic accidents causing bodily injuries and all accidents causing material damage reported by police in Quebec.

The description of the control variables can be found in Appendix 2.

5.1 Descriptive statistics

- By fleet

The data contains 17,542 fleets with a follow-up of at least two periods. On December 31st, 1998, a fleet had an average of nearly seven years of experience, with a minimum of one year and three months and a maximum of 20 years and 9 months. We note in Table 1 that approximately 4% of the 17,542 fleets have over 20 trucks. On average, a truck has 4.11 observation periods ranging from 3.38 to 5.71.

Table 1: Size of fleet distribution

Size of fleet	N	%	Number of trucks	Number of periods by truck	
				Mean	Median
2	6,888	39.27	13,376	3.38	3
3	3,203	18.26	9,609	4.07	4
4 to 5	3,285	18.73	14,397	4.18	4
6 to 9	2,171	12.38	15,364	4.30	4
10 to 20	1,298	7.40	17,506	4.30	4
21 to 50	496	2.83	15,042	4.25	4
More than 50*	197	1.12	22,355	3.92	3
Biggest fleets	4	0.01	3,057	5.71	6
Total	17,542	100.00	111,106	4.11	4

*Excluding the biggest fleets

In Table 2 we observe that a quarter of the 17,542 fleets have eight years of follow-up, which confirms the panel aspect of the data. We also note that there are 3,649 fleets for which we have two consecutive years of follow-up, which is 99.5% of fleets with two observation periods.

Table 2: Number of years of follow-up of the firm

Number of years of follow-up	Fleet		Truck		Truck of the 4 biggest fleets	
	N	%	N	%	N	%
2	3,649	20.80	30,716	27.65	294	9.62
3	2,512	14.32	23,270	20.94	334	10.93
4	2,075	11.83	17,831	16.05	379	12.40
5	1,654	9.43	11,998	10.80	295	9.65
6	1,645	9.38	9,241	8.32	403	13.18
7	1,567	8.93	6,225	5.60	365	11.61
8	4,440	25.31	11,825	10.64	997	32.61
Total	17,542	100.00	111,106	100.00	3,057	100.00

Table 3 shows the distribution of the size of the fleet by year. In 1991, there are 8,650 fleets. This number increases over time for a total of 87,771 fleet-years. Among the 87,771 fleet-years, 46.51% have two vehicles and about 3% have over 20 vehicles.

Table 3: Size of fleet distribution (in %) by year

Size of fleet	% by year								% total
	1991	1992	1993	1994	1995	1996	1997	1998	
2	47.86	46.31	46.60	46.82	45.98	46.08	45.83	47.03	46.51
3	19.63	19.75	19.92	19.71	19.45	19.30	19.28	19.10	19.51
4 to 5	15.26	15.99	15.75	15.54	16.28	16.02	16.40	16.26	15.96
6 to 9	9.16	9.40	9.19	9.46	9.62	9.88	9.62	9.27	9.47
10 to 20	5.45	5.72	5.76	5.74	5.76	5.68	5.95	5.64	5.72
21 to 50	1.97	2.06	2.00	1.89	2.05	2.14	2.08	2.01	2.03
More than 50	0.68	0.78	0.78	0.83	0.86	0.89	0.85	0.70	0.80
Number of fleets	8,650	10,691	11,132	11,445	11,733	11,965	11,834	10,321	87,771

Source: Anger et al (2018).

The average accident rate of trucks per fleet is lower for the year 1997. The years 1991, 1992 and 1995 had the highest recorded average rates of truck accidents per fleet (Table 4).

Table 4: Average truck accidents per fleet according to size of fleet and year

Size of fleet	Average truck accidents per fleet by year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
2	0.2626	0.2480	0.2219	0.2215	0.2219	0.2155	0.1809	0.2186	0.2224
3	0.4370	0.4154	0.3811	0.4007	0.4194	0.3712	0.3129	0.4049	0.3909
4 to 5	0.6689	0.6864	0.6030	0.6296	0.6408	0.5863	0.5507	0.6490	0.6239
6 to 9	1.3914	1.2259	1.0909	1.1311	1.1833	1.0981	1.0018	1.1996	1.1550
10 to 20	2.6730	2.6127	2.3744	2.5099	2.4527	2.4824	2.0767	2.5223	2.4497
21 to 50	5.9176	5.3818	5.0448	5.3565	5.8875	5.2461	4.8618	5.7681	5.4094
More than 50	22.6780	22.4096	21.7701	22.0421	22.0198	21.0935	18.4700	22.5417	21.5014
Average truck accidents by fleet	0.8575	0.8561	0.7824	0.8157	0.8531	0.8153	0.7106	0.8120	0.8109

Source: Anger et al (2018).

- By truck

There are 43,037 trucks in 1991 for a total of 456,177 truck-years with a mean annual truck accident rate of 15% (Table 5).

Table 5: Number of truck accidents distribution according to the year of observation

Truck accidents	% (by year of observation)								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
0	85.61	86.07	87.19	86.66	86.47	87.04	88.44	86.52	86.78
1	12.22	11.93	11.05	11.47	11.53	11.14	10.11	11.56	11.36
2	1.82	1.67	1.47	1.58	1.65	1.49	1.24	1.60	1.56
3	0.28	0.27	0.23	0.23	0.28	0.26	0.17	0.26	0.24
4 and more	0.07	0.06	0.05	0.06	0.07	0.06	0.04	0.06	0.06
Number of trucks	43,037	55,388	57,795	59,347	61,917	63,749	62,552	52,392	456,177
Mean accidents	0.1696	0.1632	0.1489	0.1556	0.1596	0.1515	0.1327	0.1578	0.1541

Source: Anger et al (2018).

Violations committed by drivers and fleet owners are usually very powerful in explaining truck accidents during the next year (Table 6). Indeed, we observe that year *t* accident rate is an increasing function of previous year violations committed by the drivers and fleet owners.

Table 6: Average truck accidents according to the driver's violations committed the previous year

	Year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
For speeding									
0	0.1642	0.1586	0.1432	0.1486	0.1516	0.1435	0.1240	0.1498	0.1472
1	0.2974	0.2592	0.2640	0.2723	0.2631	0.2523	0.2161	0.2609	0.2556
2	0.2701	0.3410	0.3045	0.4000	0.3566	0.3249	0.3207	0.3281	0.3337
3 and more	0.4194	0.5000	0.2424	0.4651	0.5506	0.4821	0.3973	0.4600	0.4505
For driving with a suspended license									
0	0.1696	0.1629	0.1485	0.1547	0.1584	0.1507	0.1321	0.1574	0.1535
1 and more	0.7500	0.5217	0.3750	0.4076	0.3549	0.3426	0.3017	0.3265	0.3566
For running a red light									
0	0.1679	0.1617	0.1473	0.1538	0.1571	0.1491	0.1308	0.1555	0.1521
1	0.2726	0.2846	0.2764	0.2999	0.3350	0.3135	0.2981	0.3413	0.3036
2 and more	0.5294	0.6667	0.3846	0.2727	0.6000	0.7272	0.2308	0.2727	0.5040
For stop signs or police signals									
0	0.1677	0.1618	0.1474	0.1541	0.1572	0.1498	0.1315	0.1561	0.1524
1	0.3204	0.3140	0.2797	0.2823	0.3570	0.2931	0.2411	0.3100	0.2993
2 and more	0.5000	0.2857	0.2500	0.5833	0.2941	0.5263	0.3125	0.5000	0.4016
For failing to wear a seat belt									
0	0.1689	0.1626	0.1481	0.1554	0.1588	0.1508	0.1316	0.1576	0.1534
1	0.2304	0.2246	0.2293	0.1770	0.2376	0.2100	0.2096	0.2124	0.2164
2 and more	0.4138	0.4333	0.2571	0.2750	0.1774	0.2653	0.3137	0.1200	0.2741
Violations committed by the driver the previous year									
For overweight									
0	0.1649	0.1583	0.1448	0.1517	0.1544	0.1461	0.1293	0.1540	0.1497
1	0.2430	0.2764	0.2410	0.2501	0.2432	0.2383	0.1889	0.2631	0.2394

	Year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
2 and more	0.3387	0.2956	0.3364	0.2926	0.3552	0.3026	0.2155	0.3874	0.3065
For oversize									
0	0.1695	0.1632	0.1488	0.1554	0.1596	0.1515	0.1326	0.1577	0.1540
1 and more	0.2836	0.1000	0.2917	0.2603	0.1574	0.1545	0.2269	0.2821	0.2119
For poorly secured loads									
0	0.1688	0.1625	0.1482	0.1550	0.1587	0.1509	0.1323	0.1570	0.1534
1 and more	0.3185	0.3198	0.2667	0.2665	0.2656	0.2621	0.2214	0.3778	0.2791
For failure to respect hours of service									
0	0.1696	0.1632	0.1486	0.1556	0.1592	0.1513	0.1325	0.1575	0.1539
1 and more	0.5714	0.3000	0.6333	0.1951	0.3529	0.2743	0.3881	0.3571	0.3496
For failure to undergo mechanical inspection									
0	0.1691	0.1626	0.1474	0.1546	0.1578	0.1509	0.1321	0.1572	0.1532
1 and more	0.2890	0.3180	0.2388	0.2534	0.3024	0.2251	0.2168	0.2768	0.2591

Source: Anger et al (2018).

5.2 Estimation results

We first estimate the two models with all observations. The results of the gamma/Dirichlet random effects model are presented in columns 2 and 3 of Table 7. Those from the first hierarchical random effect model with the same number of observations are presented in columns 4 and 5 of the table (Hierarchical I). To estimate the hierarchical random effect model, we truncated the infinite summation at the value K =13. We do not obtain better results with higher values of K. In order to diminish the time of convergence, we computed the first and second derivatives of the likelihood function of the hierarchical model to obtain the gradient and the hessian. We observe in Table 7 that the log likelihood values are very similar, and the estimated standard errors are very stable between the two models. Other estimations with different values of K (5, 10, 11, and 12) and different fleet sizes are presented in the Online appendix. When K increases, we observe from table O1 and O2 in the Online appendix that the standard error estimates of the explanatory variables remain fairly stable. This is not the case for all the coefficient estimates, however. We can see important differences at *K*=5 when compared to *K* values equal or greater than 10. The coefficient estimates do not vary very much for *K* at 10 or 13.

Several variables measure observable heterogeneity. Some of these variables (type of fuel, number of cylinders, number of axles, type of vehicle used) are characteristics concerning the vehicles, whereas others (sector, fleet size, etc.) are related to the fleet. We also include the number of violations of trucking standards the year before the accidents and the number of violations of the road safety code leading to demerit points the year before the accidents**.** The first group of violations is more related to fleet owner behavior while the second group is more related to driver behavior**.** Almost all coefficients of these variables are significant at 1% in Table 7.

We observe minor differences between the two models, however. The estimated coefficients corresponding to the observation period are not stable between the two random models. From the gamma/Dirichlet model, the years 1991 and 1992 are not statistically different from 1998 at 1% while, for the hierarchical model, we observe that the years 1993 and 1996 are not statistically different from 1998 at 1%, even if the estimated standard errors are very stable between the two models. With few exceptions the other coefficients that are significant at 1% in the gamma/Dirichlet model are also significant at 1% with the same sign in the hierarchical model. However, the general public trucking coefficient is not statistically different from bulk public trucking at 1% in the hierarchical model.

The random effects parameters are all significant in both models. In the gamma/Dirichlet model, the significance of the three random effects parameters means that the random effect associated with the fleets (or the non-observable risk of the fleets) ($\hat{\kappa}$), as well the random effects of the trucks including the drivers ($\hat{\nu}$) and the random time effects ($\hat{\rho}$) significantly affect the trucks distribution of accidents even when we control for many observable characteristics. In the hierarchical model, the four random effects parameters are also significant. In the hierarchical model, we assume that: i) the fleet effect (N_i) follows a Negative Binomial distribution with parameters $(\delta, \beta c)$; ii) the truck effect (Z_{ij}) given N_i is Negative Binomial with parameters $(\delta + N_i, \frac{\beta c}{1+\beta c})$ and, iii) the time effect (θ_{ijt}) given Z_{ij} follows a gamma distribution with parameters $(\delta^* + \beta^* Z_{ij}, c^*)$. Because all these coefficients are significant, we do not reject our assumptions. The pricing formula of the hierarchical model will have to account for this additional information.

We can compare the results of these two models with the Hausman model. The estimation results for firms with two trucks or more using the Hausman model are presented in columns 2 and 3 of Table 7b in Appendix 3 along with the results of the gamma/Dirichlet and Hierarchical I models of Table 7. The Hausman model is suitable for estimating parameters with individual effects but cannot take into account the common or the fleet effect when individual observations belong to different firms with common characteristics that can affect accident distributions. Almost all coefficients of the Hausman model are significant at 1% with the exception of the year's variables. For models estimated from the same data set, we can use the BIC and the AIC for comparison. We observe in Table 8 that the gamma-Dirichlet model performs better than the Hausman model. We also obtain better estimation results with the Hierarchical I model (K=13) than with the gamma/Dirichlet model according to the different criteria presented in Table 8.

We also estimated the hierarchical model with fewer observations and with K=13 (Hierarchical II in Table 7). When a fleet has many trucks, the value of the multiplicative terms on all trucks of a fleet in equation (8) becomes too small (near zero) for estimation, so we had to bound the value of log $M_i(X)$ to log (10^{-300}) for the 4 biggest fleets to optimize the log of the likelihood function corresponding to the Hierarchical I model. This procedure may lead to a bias. We then estimated the hierarchical model by simply dropping the observations of the four biggest fleets and with K=13.

The new results are presented in Table 7. The Hierarchical II model has been optimized with 438,717 observations while the Hierarchical I model has been optimized with 456,177 observations. The difference is obtained by dropping 3,057 trucks belonging to the 4 biggest fleets resulting in 108,049 trucks. We see from Table 7 that removing the trucks from the four biggest fleets does not influence the standard errors estimates but affects some coefficients. The biggest difference is in the value of $\hat{\beta}^*$ going from 10 to 4. From the discussion below equation (2), we observe that $\hat{\beta}^*$ affects the correlation of the temporal random effects for a given truck in a given fleet. A value of 4 seems more reasonable than a value of 10. There are also differences for some control variables such as the sector of activity of the fleet, type of fuel, number of axles, and observation periods. The log likelihood value improved as well. The coefficient estimates of the hierarchical Model II are more similar to those from the gamma/Dirichlet Model except for the number of axles.

Table 7: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991-1998 period (fleet of two trucks or more and trucks with two periods or more), with the gamma/Dirichlet model and two hierarchical random effect models with K = 13.

Explanatory variable	Gamma/Dirichlet		Hierarchical I		Hierarchical II	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.9070*	0.0573	-3.6435*	0.0548	-3.3297*	0.0528
Number of years as a fleet	-0.0464*	0.0044	-0.0260*	0.0040	-0.0459*	0.0039
Sector of activity in 1998						
Other sector	-0.1426	0.1163	-0.2681	0.1118	-0.2784	0.1272
General public trucking	0.1685*	0.0304	0.0727	0.0324	0.1536*	0.0314
Bulk public trucking	Reference group		Reference group		Reference group	
Private trucking	0.2290*	0.0256	0.2007*	0.0269	0.1232*	0.0247
Short-term rental firm	0.5633*	0.0483	0.1391*	0.0450	0.4998*	0.0495
Size of fleet						
2	Reference group		Reference group		Reference group	
3	0.0801*	0.0205	0.1287*	0.0194	0.1192*	0.0187
4 to 5	0.1385*	0.0205	0.2104*	0.0196	0.1758*	0.0185
6 to 9	0.2137*	0.0210	0.3114*	0.0209	0.2366*	0.0194
10 to 20	0.2937*	0.0209	0.4510*	0.0221	0.2790*	0.0201
21 to 50	0.3010*	0.0223	0.6678*	0.0258	0.2305*	0.0235
More than 50	0.3077*	0.0217	1.5852*	0.0261	0.2412*	0.0308
Days in previous year	2.0537*	0.0300	1.7784*	0.0298	1.6821*	0.0297
Violations						
For overload	0.0966*	0.0115	0.0809*	0.0119	0.0977*	0.0119
For excessive size	0.1480	0.0860	0.1448	0.0884	0.1550	0.0890
For poorly secured cargo	0.2054*	0.0354	0.1826*	0.0365	0.2032*	0.0368
Not respect service hours	0.1984*	0.0664	0.1984*	0.0678	0.2003*	0.0681
No mechanical inspection	0.1778*	0.0298	0.1575*	0.0307	0.1893*	0.0307
For other reasons	0.1754	0.0743	0.2113*	0.0771	0.2222*	0.0770
Type of vehicle use						
Commercial use	-0.1938*	0.0212	-0.1443*	0.0231	-0.1789*	0.0221
Other than bulk goods	-0.1148*	0.0243	-0.1159*	0.0268	-0.0800*	0.0258
Bulk goods	Reference group		Reference group		Reference group	
Type of fuel						
Diesel	Reference group		Reference group		Reference group	
Gas	-0.3973*	0.0136	-0.3441*	0.0153	-0.3819*	0.0149
Other	-0.3079*	0.0736	-0.4090*	0.0824	-0.2886*	0.0818
Number of cylinders						
1 to 5	0.2167*	0.0403	0.3656*	0.0462	0.3149*	0.0452
6 to 7	0.3780*	0.0126	0.3540*	0.0143	0.3308*	0.0139
8 or more than 10	Reference group		Reference group		Reference group	

Explanatory variable	Gamma/Dirichlet		Hierarchical I		Hierarchical II	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Number of axles						
2 axles (3,000 to 4,000 kg)	-0.2916*	0.0208	-0.2898*	0.0233	-0.3049*	0.0230
2 axles (4,000 kg and more)	-0.2850*	0.0150	-0.2856*	0.0173	-0.3140*	0.0167
3 axles	-0.1278*	0.0149	-0.1641*	0.0170	-0.2828*	0.0165
4 axles	-0.1321*	0.0190	-0.1590*	0.0222	-0.1951*	0.0214
5 axles	-0.1973*	0.0174	-0.1914*	0.0194	-0.2506*	0.0189
6 axles or more	Reference group		Reference group		Reference group	
Number of violations						
For speeding	0.1946*	0.0103	0.1849*	0.0107	0.2152*	0.0107
Suspended license	0.3830*	0.0422	0.3740*	0.0430	0.3876*	0.0432
For running a red light	0.3094*	0.0239	0.2815*	0.0246	0.3151*	0.0249
For ignoring a stop sign	0.3597*	0.0258	0.3150*	0.0266	0.3535*	0.0267
Not wearing a seat belt	0.1568*	0.0294	0.1362*	0.0303	0.1741*	0.0303
Observation period						
1991	0.0760	0.0332	0.1990*	0.0308	0.0023	0.0299
1992	0.0548	0.0293	0.1085*	0.0272	-0.0379	0.0265
1993	0.0806*	0.0259	0.0223	0.0244	-0.1039*	0.0239
1994	0.1845*	0.0226	0.0569*	0.0215	-0.0344	0.0212
1995	0.2073*	0.0197	0.0581*	0.0190	-0.0088	0.0189
1996	0.1198*	0.0175	-0.0201	0.0171	-0.0449*	0.0172
1997	-0.0791*	0.0163	-0.1613*	0.0163	-0.1547*	0.0166
1998	Reference group		Reference group		Reference group	
$\hat{\delta}$			0.8168*	0.0250	0.5985*	0.0309
$\widehat{\beta c_0}$			3.0504*	0.0912	1.8361*	0.0599
$\hat{\delta}^*$			4.7817*	1.4580	4.0598*	0.6267
$\hat{\beta}^*$			10.1139*	3.1455	4.2851*	0.6934
$\hat{\nu}$	2.0086*	0.0422				
$\hat{\kappa}$	12.6597*	0.2508				
$\hat{\rho}$	4.6670*	0.3102				
Number of observations	456,177		456,177		438,717	
Number of trucks	111,106		111,106		108,120	
Log likelihood	-197,116		-193,026		-186,275	
$\widehat{\beta c} = \frac{\widehat{\beta c_0}}{1 + \widehat{\beta c_0}}$			0.7531		0.6474	

* Significant at 1%.

Table 8: Fit statistics of the three models with all observations log likelihood

Statistics	Hausman model	Gamma-Dirichlet model	Hierarchical I model K=13
Log likelihood	-197,165	-197,116	-193,036
BIC	394,904	394,819	386,651
AIC	394,418	394,322	386,144
Number of trucks	111,106	111,106	111,106
Number of observations	456,177	456,177	456,177
Number of parameters	44	45	46

Note: Bayesian Information Criterion (BIC) = –2lnL + kln(N); Akaikes Information Criterion (AIC) = –2lnL + 2k, where *k* and *N* are the number of parameters and observations respectively. The Likelihood ratio test value (8,160) is largely superior to the critical value of 6.63 at 1% when comparing the gamma-Dirichlet model to the Hierarchical I model. The Likelihood ratio test value of 98 is also largely superior to the same critical value when comparing the Hausman model to the gamma-Dirichlet model.

6. Empirical pricing model

6.1 Empirical pricing formula

We can use the estimated parametric models to rate the insurance for vehicles belonging to a fleet. According to the results in Table 7, a premium will have to be a function of observable characteristics of the vehicle and the fleet as well as a function of violations of the Highway Safety Code committed by drivers and fleets, two variables that approximate the asymmetric information between the insurer and both the fleet owners and drivers. As previously stated, this will not be enough, however, to obtain accurate pricing because many unobservable characteristics of trucks, drivers and fleets also affect the trucks' distribution of accidents. The premiums will have to be adjusted using the parameters of the random effects in order to account for the impact of the unobservable characteristics of fleets and trucks as well as owners and drivers' behaviors, and even time that is not captured by year variables. This form of rating makes it possible to visualize the impact (observable and non-observable) of the behaviors of owners and drivers on the predicted rate of accidents, and consequently on premiums under potential asymmetric information between the insurer and the trucking firm. We now present the empirical pricing formula of the model.

Our goal is to build a bonus-malus system based on the number of past accidents and control variables in the regression model. We use the expected value principle for the premium of a truck in a given fleet. With the hierarchical model, to construct an optimal bonus-malus scheme based on the number of past accidents recorded for a truck in a given fleet as well as those observed for all trucks of its fleet during the same period, we calculate the posterior expected number of accidents at period *T*+1 for a truck *j* of a given fleet *i*:

$$\begin{aligned} \mathrm{E}\left[X_{i,j,T+1}|x_{i,k,t}\right] &= \lambda_{i,j,T+1}\mathrm{E}\left[\theta_{i,j,T+1}|x_{i,k,t}\right] \\ &= \lambda_{i,j,T+1}\mathrm{E}\left[c^*\left(\delta^*+\beta^* Z_{i,j}\right)|x_{i,k,t}\right] \\ &= \lambda_{i,j,T+1}c^*\delta^*+\lambda_{i,j,T+1}\beta^* c^*\mathrm{E}\left[Z_{i,j}|x_{i,k,t}\right] \end{aligned} \tag{16}$$

where $\lambda_{i,j,T+1}$ is the marginal expectation of $X_{i,j,T+1}$ given all observable covariates. The conditional expectation $\mathrm{E}\left[Z_{i,j}|x_{i,k,t}\right]=\sum_{z_j=0}^{\infty} z_j\, p\left(z_j|x_{i,k,t}\right)$ and $p\left(z_j|x_{i,k,t}\right)$ is given by equation (13).

In order to evaluate how the estimated parameter differences of the random effects may influence the value of the posterior expected number of accidents at period *T*+1, we calculate its value with two sets of parameters. The first set uses the parameters estimated with the Hierarchical I model in Table 7, which are equal to $\hat{\delta}$ = 0.8168; $\widehat{\delta^*}$ = 4.7817; $\widehat{\beta^*}$ = 10.1139; $\widehat{\beta c}$ = 0.7531 and $\widehat{c^*} = 1\Big/\left(\widehat{\delta^*}+\left[\hat{\delta}\widehat{\beta^*}\frac{\widehat{\beta c}}{1-\widehat{\beta c}}\right]\right)$ = 0.0334. The estimated $\lambda_{i,j,T+1}$, is the mean over *t* of the optimal estimated $\lambda_{i,j,t}$, which are the exponential of the sum of the control variables' estimated parameters in Table 7 multiplied by their corresponding control variable values of truck *j* in fleet *i* at period *t*. The second set uses the parameters estimated with the Hierarchical II model of table 7, which are equal to $\hat{\delta}$ = 0.5985; $\widehat{\delta^*}$ = 4.0598; $\widehat{\beta^*}$ = 4.2851; $\widehat{\beta c}$ = 0.6474 and $\widehat{c^*} = 1\Big/\left(\widehat{\delta^*}+\left[\hat{\delta}\widehat{\beta^*}\frac{\widehat{\beta c}}{1-\widehat{\beta c}}\right]\right)$ = 0.1140.

We can see from Figure 2 that the distribution of the posterior expected number of accidents at period *T*+1 is more dispersed for the Hierarchical I model (at bottom) than for the Hierarchical II model (at top). The corresponding means and standard deviations (in parentheses) are respectively equal to 0.1619 (0.1271) and 0.1531 (0.1100). Both distributions differ significantly from the normal distribution.

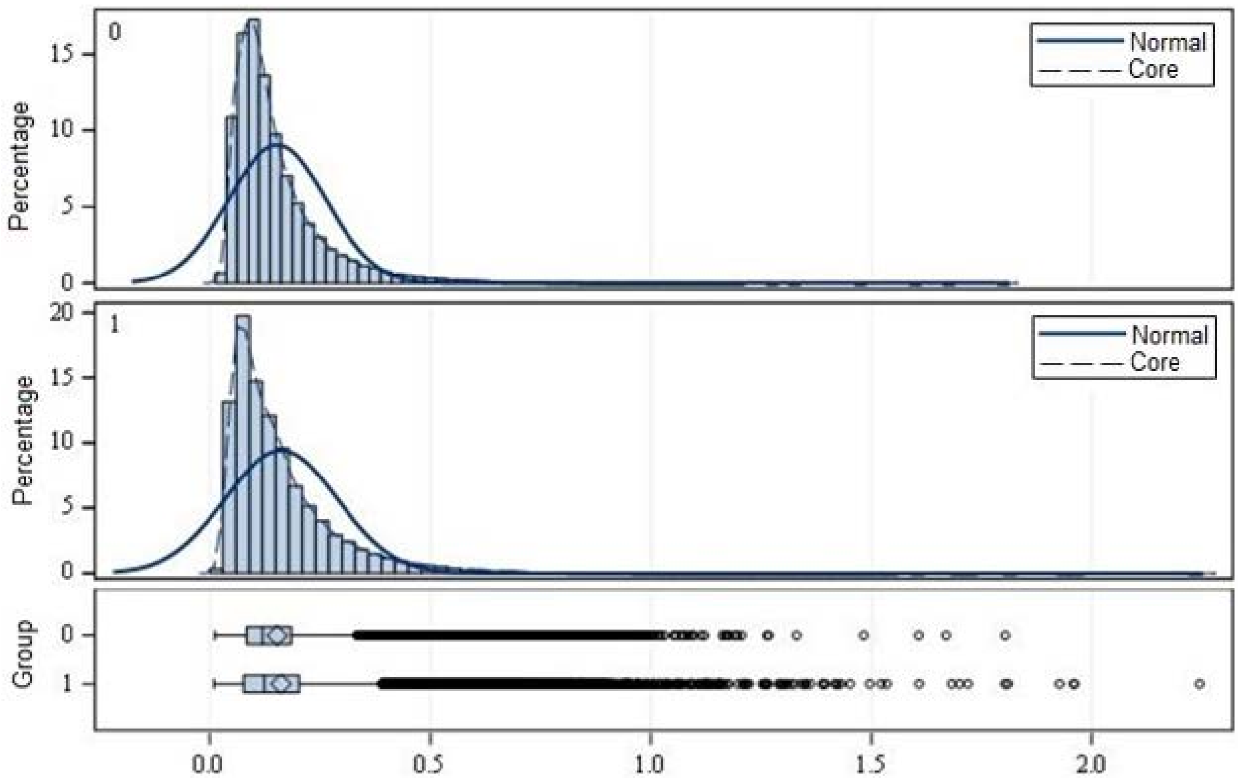

Figure 2: Posterior expected number of accidents at period T+1 obtained from the estimation of Hierarchical II model (at top) and Hierarchical I model (at bottom)

To compare the two values of the posterior expected number of accidents at time T+1, we perform a paired t-test of the difference of the posterior expected number of accidents at period T+1. The mean of the difference is statistically different from zero at 1%: its value is equal to 0.00879 and the t-test value is 73.48. The mean value of the difference is small however. Figure 3 presents the distribution of the difference used for the paired t-test. It also differs from the normal distribution.

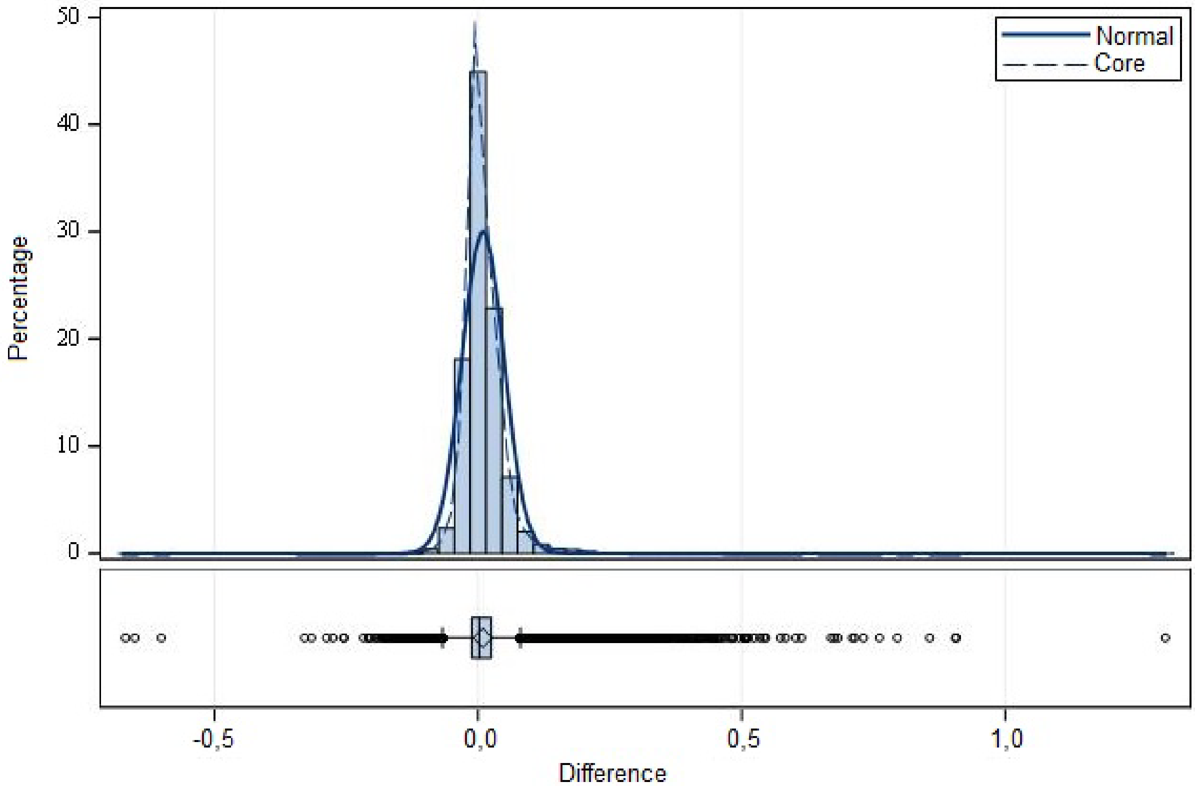

Figure 3: Distribution of the difference between the Hierarchical I model and the Hierarchical II model of the posterior expected number of accidents at period T+1.

6.2 Out-of-sample validation

To test for the forecasting performance of the models, we perform out-of-sample validations. For the test, the data for estimation are from the observation period 1991-1997. The model is then tested on the data from the validation year 1998. This a very severe test because the validation period is very short.

The database for the re-optimization of the Hierarchical I model contains 393,634 observations. We also re-estimated the Hierarchical II model with 378,113 observations obtained by dropping the 2,950 trucks related to the four biggest fleets. The results are presented respectively in the Hierarchical III model and the Hierarchical IV model in Table 9. We still observe a big difference between the two estimated $\hat{\beta}^*$ and model IV has a higher log likelihood value than model III. The other parameters are stable.

Table 9: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991-1997 period (fleets of two trucks or more) and trucks with two periods or more. Hierarchical random effect models with K = 13.

Explanatory variable	Hierarchical III model		Hierarchical IV model	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.7866*	0.0586	-3.4438*	0.0574
Number of years as a fleet	-0.0272*	0.0047	-0.0499*	0.0048
Sector of activity in 1998				
Other sector	-0.2515	0.1259	-0.2802	0.1351
General public trucking	0.0759	0.0351	0.1876*	0.0331
Bulk public trucking	Reference group		Reference group	
Private trucking	0.2146*	0.0291	0.1366*	0.0264
Short-term rental firm	0.1549*	0.0507	0.4860*	0.0534
Size of fleet				
2	Reference group		Reference group	
3	0.1222*	0.0207	0.1129*	0.0199
4 to 5	0.2061*	0.0210	0.1733*	0.0197
6 to 9	0.2912*	0.0223	0.2206*	0.0207
10 to 20	0.4372*	0.0238	0.2679*	0.0215
21 to 50	0.6429*	0.0284	0.2191*	0.0254
More than 50	1.6176*	0.0334	0.2574*	0.0361
Days in previous year	1.7849*	0.0319	1.6803*	0.0318
Violations				
For overload	0.0828*	0.0125	0.0984*	0.0124
For excessive size	0.1188	0.0926	0.1272	0.0935
For poorly secured cargo	0.16667*	0.0388	0.1851*	0.0390
Not respect service hours	0.2018*	0.0775	0.2158*	0.0783
No mechanical inspection	0.1468*	0.0323	0.1788*	0.0323
For other reasons	0.1997	0.0845	0.2140	0.0844
Type of vehicle use				
Commercial use	-0.1527*	0.0249	-0.1866*	0.0236
Other than bulk goods	-0.1185*	0.0288	-0.0717*	0.0275
Bulk goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.3317*	0.0160	-0.3720*	0.0155
Other	-0.4592*	0.0903	-0.3447*	0.0893
Number of cylinders				
1 to 5	0.3757*	0.0495	0.3221*	0.0485
6 to 7	0.3521*	0.0151	0.3286*	0.0146
8 or more than 10	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.2759*	0.0495	-0.2838*	0.0251
2 axles (more than 4,000 kg)	-0.2824*	0.0186	-0.3072*	0.0178
3 axles	-0.1685*	0.0184	-0.2900*	0.0177
4 axles	-0.1543*	0.0237	-0.1979*	0.0228
5 axles	-0.1889*	0.0209	-0.2466*	0.0201
6 axles or more	Reference group		Reference group	

Explanatory variable	Hierarchical III model		Hierarchical IV model	
	Coefficient	Standard error	Coefficient	Standard error
Number of violations				
For speeding	0.1892*	0.0117	0.2214*	0.0117
Suspended license	0.3779*	0.0453	0.3900*	0.0455
For running a red light	0.2742*	0.0264	0.3115*	0.0267
For ignoring a stop sign	0.3029*	0.0286	0.3433*	0.0288
Not wearing a seat belt	0.1513*	0.0315	0.1877*	0.0315
Observation period				
1991	0.3290*	0.0313	0.1275*	0.0315
1992	0.2407*	0.0272	0.0905*	0.0274
1993	0.1556*	0.0239	0.0281	0.0241
1994	0.1896*	0.0207	0.1023*	0.0209
1995	0.1920*	0.0180	0.1332*	0.0183
1996	0.1136*	0.0163	0.1012*	0.0167
1997	Reference group		Reference group	
$\hat{\delta}$	0.8326*	0.0275	0.5849*	0.0332
$\widehat{\beta c}_0$	2.9818*	0.0964	1.8252*	0.0643
$\hat{\delta}^*$	4.9122*	1.7273	4.0960*	0.5874
$\hat{\beta}^*$	11.0520*	3.9615	4.0960*	0.6533
Number of observations:	393,634		378,113	
Number of trucks	100,955		98,005	
Log likelihood	-168,013		-161,680	
$\widehat{\beta c} = \dfrac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.7489		0.6460	

* Significant at 1%.

To compare the means of the posterior expected number of accidents at period *T*+1 from estimations of the Hierarchical III and IV models in Table 9 and the number of accidents in 1998, we run a *t*-test. The results are presented in Table 10. For all fleets, we observe that the means are different at 1% with the Hierarchical model III. The mean of the posterior expected number of accidents at period *T*+1 is not statistically different from the mean of observed accidents in 1998 at 1% level of significance for the Hierarchical IV model. So the Hierarchical IV model seems to better perform than the Hierarchical III model. The differences in the means are not statistically different with the Hierarchical IV model for fleets with two trucks, fleets having 2, 20 to 50 trucks and fleets with more than 50 trucks (results not presented).

Table 10: *t*-test of the posterior expected number of accidents at period *T*+1 from estimations of Hierarchical III and IV models and the observed numbers of accidents in 1998 for all fleets.

		Hierarchical III model			Data 1998		*t*-test	
	N	Mean	Std	N	Mean	Std	*t*-value	*p*-value
All fleets	100,955	0.1631	0.1282	52,392	0.1578	0.4332	2.73	<0.0001
		Hierarchical IV model			Data 1998		*t*-test	
	N	Mean	Std	N	Mean	Std	*t*-value	*p*-value
All fleets	98,005	0.1533	0.1093	50,558	0.1575	0.4317	-2.11	0.035

Mean: Posterior expected number of accidents at period *T*+1 from estimations in Table 9
Data 1998: Number of accidents in 1998

From the results given at Table 10, the posterior expected numbers of accidents from estimations of Model IV in Table 9 seem to better estimate the posterior expected number of accidents at period *T*+1. In the next section, we present premium tables derived from the posterior expected number of accidents at period *T*+1. From the out-of-sample results, we propose using the estimated parameters of Hierarchical II model in Table 7 for the premiums. This model corresponds to Hierarchical IV model in the out-of-sample test. For comparison, we will also present those obtained with the estimated parameters of Hierarchical I model in Table 7, even if this model does not perform very well out-of-sample.

6.3 Application of the bonus-malus system

In this section, we propose premium tables. Given that we did not have data to compute the conditional average cost of claims, we use \$10,000, a seemingly reasonable value for accidents involving trucks in North America during that period (Dionne et al, 1999).

A premium for a truck at period *T*+1 has three possibilities:

1. It has past experience belonging to a fleet having trucks with past experience;
2. It is a new truck in a new fleet, meaning no past experience for the truck and no past experience for the fleet;
3. It is a new truck belonging to an existing fleet that has past experience.

We now consider these three possibilities.

Premiums for a truck with past experience belonging to a fleet with past experience

Table 11 presents premiums for a truck using all information from the optimal estimations of the Hierarchical II model in Table 7 to be coherent with the out-of-sample results presented in Table 10. From the Hierarchical II model and equation (16), the average estimated number of accidents in period *T*+1 is 0.1538. This is very close to the empirical mean of 0.1541presented in Table 5. It decreases to 0.1056 if the truck did not accumulate accidents over the past but increases to 0.3487 if it has accumulated three accidents in the past. The variations are similar for the Hierarchical I model.

Table 11: Premiums for a truck as a function of accumulated number of past accidents

		Hierarchical I model		Hierarchical II model	
	N	$E[X_{ijT+1}\|X_{ikt}]$	Premium	$E[X_{ijT+1}\|X_{ikt}]$	Premium
Accumulated number of accidents over the period					
0	69,219	0.1008	$1,008	0.1056	$1,056
1	25,772	0.1976	$1,976	0.1750	$1,750
2	9,498	0.2995	$2,995	0.2595	$2,595
3	3,765	0.3944	$3,944	0.3487	$3,487
More than 3	2,852	0.5583	$5,583	0.4961	$4,961
	111,106	0.1631	$1,631	0.1538	$1,538

Posterior expected number of accidents in period *T*+1 from estimations in Table 7, equation (16) and conditional average cost of claims of $10,000.

Premiums for a new truck belonging to a new fleet

If the fleet does not exist in periods 1 to *T*, then there are no past accidents recorded for the new truck. In this case $\mathrm{E}\left[\theta_{i,j,T+1} | x_{i,k,t}\right] = \mathrm{E}\left[\theta_{i,j,t}\right] = 1$.

The posterior expected number of accidents at period *T*+1 for a new truck of a new fleet will be:

$$\mathrm{E}\left[X_{i,j,T+1} | x_{i,k,t}\right] = \lambda \qquad (17)$$

where λ is the mean over *i* of the optimal estimated λ_i, λ_i is the mean over *j* of the optimal estimated λ_{ij} and λ_{ij} is the mean over *t* of the optimal estimated λ_{ijt}. Since we have no information concerning the past of the new truck and the past of a new fleet, we represent the truck by a representative vehicule of the population, i.e. λ. However, even if this is a new truck, we may know its characteristics, such as the type of gasoline it uses, the number of axles, the number of cylinders, etc. We can thus obtain various scenarios for λ, according to the characteristics of the truck.

Table 12 presents the premium for a new truck from a new fleet. Its average estimated number of accidents at period *T*+1 is 0.1471, which is lower than the value 0.1538 obtained in Table 11. For comparison we also present the values by fleet size. The premium falls to 0.1103 if the truck belongs to a new fleet of size 2. Note that the out-of-sample test did not reject this size of fleet.

Table 12: Premiums for a new truck in a new fleet

A new truck if the fleet did not exist	N	Hierarchical I model λ	Hierarchical I model Premium	Hierarchical II model λ	Hierarchical II model Premium
Size of fleet					
2	13,776	0.1207	$1,207	0.1103	$1,103
3	9,609	0.1331	$1,331	0.1200	$1,200
4 to 5	14,397	0.1469	$1,469	0.1307	$1,307
6 to 9	15,364	0.1684	$1,684	0.1465	$1,465
10 to 20	17,506	0.1994	$1,994	0.1627	$1,627
21 to 50	15.042	0.2422	$2,422	0.1693	$1,693
More than 50	25,412	0.5476	$5,475	0.1629	$1,629
	111,106	0.2582	$2,582	0.1471	$1,471

Posterior expected number of accidents in period *T*+1 from estimations in Table 7, equation (17) and conditional average cost of claims of $10,000.

Premiums for a new truck belonging to a fleet with past experience

If the fleet existed, there are no past accidents recorded for the new truck but there are past accidents observed for all other trucks of its new fleet. The posterior expected number of accidents at period *T*+1 for a new truck of a given fleet with past experience becomes:

$$\mathrm{E}\left[X_{i,j,T+1}|x_{i,k,t}\right] = \lambda_i c^* \delta^* + \lambda_{i,} \beta^* c^* \mathrm{E}\left[Z_{i,j}|x_{i,k,t}\right] \quad (18)$$

where λ_i is the mean over j of the optimal estimated $\lambda_{i,j}$ and $\mathrm{E}\left[Z_{i,j}|x_{i,k,t}\right]$ is the average over j of the conditional expectation $\mathrm{E}\left[Z_{i,j}|x_{i,k,t}\right]$. Table 13 presents the premiums for a new truck in an existing fleet. The estimated number of accidents at period T+1 is 0.1536, a number very similar to that obtained in Table 11 but it falls to 0.1078 instead of 0.1056 if the truck belongs to a fleet of size 2.

Table 13: Premiums for a new truck in an existing fleet

A new truck if the fleet existed		Hierarchical I model		Hierarchical II model	
	N	$E[X_{ijT+1}\|X_{ikt}]$	Premium	$E[X_{ijT+1}\|X_{ikt}]$	Premium
Size of fleet					
2	13,776	0.1075	$1,075	0.1078	$1,078
3	9,609	0.1187	$1,187	0.1187	$1,187
4 to 5	14,397	0.1322	$1,322	0.1316	$1,316
6 to 9	15,364	0.1498	$1,498	0.1489	$1,489
10 to 20	17,506	0.1741	$1,741	0.1722	$1,722
21 to 50	15.042	0.1922	$1,922	0.1879	$1,879
More than 50	25,412	0.2273	$2,273	0.1779	$1,779
	111,106	0.1981	$1,981	0.1536	$1,536

Posterior expected number of accidents in period T+1 from estimations in Table 7, equation (18) and conditional average cost of claims of $10,000.

7. Conclusion

In this paper we derive a new hierarchical random effect model for the estimation and forecasting of truck accidents as an alternative to the gamma-Dirichlet approach recently introduced in the literature by Angers et al (2018). We derive a closed form formula for the ratemaking of a bonus-malus scheme that considers past accidents, past road safety offences of both drivers and fleet owners and characteristics of trucks and fleets including their size.

The estimation results of the model dominate those of the previous models in the literature including the Hausman model, which is restricted to individual random effects. However, the estimation results and out-of-sample tests of the new model are better when we drop the largest fleets in the data. This may be explained by the fact that trucks in very large firms are present in the panel in more periods than those in small fleets, as documented in Table 1. One possible extension would be to reconsider the model for all possible fleet sizes by using fixed effects instead of random effects, which is beyond the scope of this article.

References

Angers, J.F., Desjardins, D., Dionne, G., and Guertin, F. (2018). Modelling and estimating individual and firm effects with count panel data. *ASTIN Bulletin* 48(3), 1049–1078.

Angers, J.F., Desjardins, D., Dionne, G., and Guertin, F. (2006). Vehicle and fleet random effects in a model of insurance rating for fleets of vehicles. *ASTIN Bulletin* 36(1), 25–77.

Antonio, K., Frees, E.W., and Valdez, E.A. (2010). A Multilevel analysis of intercompany claim counts. *ASTIN Bulletin* 40(1), 151–177.

Baltagi B.H. (1995). *Econometric Analysis of Panel Data.* Wiley, Chichester.

Boucher J.P., Denuit M. (2006). Fixed versus random effects in Poisson regression models for claim counts: A case study with motor insurance. *Astin Bulletin* 36, 285-301.

Boucher J.P, Denuit, M., Guillen, M. (2008). Models of insurance claim counts with time dependence based on generalization of Poisson and negative binomial distributions. *Casualty Actuary Society* 2, 135-162.

Boucher, J., Guillén, M. (2009). A survey on models for panel count data with applications to insurance. Rev. R. Acad. Cien. Serie A. Mat. 103, 277-294. https://doi.org/10.1007/BF03191908.

Cameron, A.C., Trivedi, P.K. (1986). Econometric models based on count data: Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics* 1, 29–54.

Cameron A.C., Trivedi P.K. (2013). *Regression Analysis of Count Data*, Second Edition, Cambridge University Press.

Desjardins, D., Dionne, G., Pinquet, J. (2001). Experience rating schemes for fleets of vehicles. *ASTIN Bulletin* 31(1), 81-106.

Dionne, G., Laberge-Nadeau, C., Desjardins, D., Messier, S., Maag, U. (1999). Analysis of the economic impact of medical and optometric driving standards on costs incurred by trucking firms and on the social costs of traffic accidents. In: *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, G. Dionne and C. Laberge-Nadeau (Eds.), 323-351.

Dionne, G., Gagné, R., Gagnon, F., Vanasse, C. (1997). Debt, moral hazard and airline safety: An empirical evidence. *Journal of Econometrics* 79, 379-402.

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

Dionne G., Vanasse, C. (1989). A generalization of automobile insurance rating models: The negative binomial distribution with a regression component. *ASTIN Bulletin* 19, 199-212.

Fardilha, T., de Lourdes Centeno, M., Esteves, R. (2016). Tariff systems for fleets of vehicles: A study on the portfolio of Fidelidade. *European Actuarial Journal* 6, 331-349.

Frangos N., Vrontos, S.D. (2001). Design of optimal bonus-malus systems with a frequency and a security component on an individual basis in automobile insurance. *ASTIN Bulletin* 31, 1-22.

Frees, E.W., Valdez, E., 2011. Hierarchical insurance claims modeling. *Journal of the American Statistical Association* 103, 1457-1469.

Gouriéroux, C. and Lu, Y. (2019). Negative binomial autoregressive process with stochastic intensity. *Journal of Time Series Analysis* 40(2), 225–247.

Gouriéroux C., Monfort A., and Trognon A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models. *Econometrica* 52, 701-720.

Hausman J.A., Hall, B.H., and Griliches, Z. (1984). Econometric models for count data with an application to the patents – R&D relationship. *Econometrica* 52, 909-938.

Hausman J.A., Wise, D.A. (1979). Attrition bias in experimental and panel data: The Gary income maintenance experiment. *Econometrica* 47, 455-473.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell Journal of Economics* 13, 324-340.

Hsiao C. (1986). Analysis of panel data. *Econometric Society Monographs,* no 11, Cambridge University Press, Cambridge.

Laffont, J.J., Martimort, D. (2001). *The theory of incentives: The principal-agent model*. Princeton University Press, Princeton.

Norberg, R. (1986). Hierarchical credibility: analysis of a random effect linear model with nested classification. *Scandinavian Actuarial Journal* 3-4, 204–222.

Pinquet J. (2020). Positivity properties of the ARFIMA (0, d, 0) specifications and credibility analysis of frequency risks. *Insurance: Mathematics and Economics* 95, 159–165.

Pinquet J. (2013). Experience rating in non-life insurance. In Dionne, G. (Ed.), *Handbook of Insurance*. Springer, New York, 471-586.

Pitt, M.K. and Walker, S. G. (2005). Constructing stationary time series models using auxiliary variables with applications. *Journal of the American Statistical Association* 100(470), 554–564.

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

Appendix 1 – The Poisson-gamma conjugacy

The following proposition reviews the Poisson-gamma conjugacy.

Proposition 2 (Gouriéroux and Lu, 2019)

Let us consider a couple (X,Y), where X is a count variable and Y a real positive variable with joint density (with respect to $\nu \otimes \lambda +$ *, i.e., the product measure between the counting measure v on N and the Lebesgue measure* λ^+ *on* R^+*):*

$$f(x,y) = \frac{exp\left[-y\left(\beta + \frac{1-\beta c}{c}\right)\right]}{x!\Gamma(\delta)} y^{x+\delta-1}\beta^x \left(\frac{1-\beta c}{c}\right)^\delta \quad \text{(A1)}$$

with β, c *positive and* $\beta c < 1$*. Then:*

- *the conditional distribution of X given Y = y is Poisson:* $P(\beta y)$*;*
- *the conditional distribution of Y given* $X = x$ *is* : $\gamma(\delta + x, c)$ *with scale parameter c and shape parameter* $\delta + x$*;*
- *the marginal distribution of X is:* $NB(\delta, \beta c)$*;*
- *the marginal distribution of Y is:* $\gamma\left(\delta, \frac{c}{1-\beta c}\right)$.

Appendix 2 – Description of variables

The unit of observation is an eligible vehicle that is authorized to drive at least one day in year *t* and that has had a follow-up for at least two years. We analyze the accident totals found in the SAAQ files. These totals include all the traffic accidents causing bodily injuries and all accidents causing material damage reported by police in Quebec. The names of the variables in the tables of the paper are in bold at the end of description.

Dependent variable

Y_{fit} = the number of accidents in which vehicle *i* of fleet *f* has been involved during year *t*. Y_{fit} can take the values 0, 1, 2, 3, 4 and over.

Explanatory variables

We have two types of explanatory variables: those concerning the fleet and those concerning the vehicle.

Variables concerning the fleet

- *Size of fleet for year t:* 7 dichotomous variables have been created. **(Size of fleet)**

  The two-vehicle size is used as the reference category. Coefficients estimated as positive and significant will thus indicate that vehicles are more at risk of accidents than those in the two-vehicle category.

- *Sector of economic activity*: 5 dichotomous variables have been created for vehicles transporting goods:

  sect_14 = 1 if the main sector of activity is transporting passengers; **(Other sector)**
  sect_05 = 1 if the sector of activity is general public trucking; **(General public trucking)**
  sect_06 = 1 if the sector of activity is public bulk trucking; **(Bulk public trucking)**
  sect_07 = 1 if the sector of activity is independent trucking; **(Private trucking)**
  sect_08 = 1 if the sector of activity is a short-term leasing firm. **(Short-term rental firm)**

  The "public bulk trucking" sector is used as the reference category. Coefficients estimated as negative and significant for the fleet's other sectors of economic activity will thus indicate that the vehicles of these sectors run lower risks than those in the reference group (and inversely for positive and significant coefficients).

- Six variables have been created for vehicles engaged in the *transportation of goods*, so as to measure the number of convictions per vehicle in the year preceding year *t* for each fleet:

  - *Number of overweight violations per vehicle committed by a fleet in the year preceding year t.* A positive sign is predicted because more overweight violations should, on average, generate more accidents. **(For overload)**

- *Number of oversize violations per vehicle committed by a fleet in the year preceding year t:* A positive sign is predicted because more violations for oversize should, on average, generate more accidents. **(For excessive size)**
- *Number of violations per vehicle for poorly secured loads committed by a fleet in the year preceding year t:* A positive sign is predicted because more violations for poorly secured loads should, on average, generate more accidents. **(For poorly secured cargo)**
- *Number of violations per vehicle concerning hours-of-service regulations committed by a fleet in the year preceding year t:* A positive sign is predicted because more violations of hours-of-service regulations should, on average, generate more accidents. **(Not respect service hours)**
- *Number of violations per vehicle of Highway Safety Code provisions regarding mechanical inspections committed by a fleet in the year preceding year t:* A positive sign is predicted because more violations against regulations regarding mechanical inspection should, on average, generate more accidents. **(No mechanical inspection)**
- *Number of violations per vehicle, other than those already mentioned, committed by a fleet in the year preceding year t:* A positive sign is predicted because more violations other than those already mentioned should, on average, generate more accidents. (**For other reasons)**

Variables concerning the vehicle

- *Vehicle's number of cylinders***:** 4 dichotomous variables have been created:
  cyl_0 = 1 if the vehicle's number of cylinders is not known;
  cyl1_5 = 1 if the vehicle has 1 to 5 cylinders; **(1 to 5)**
  cyl6_7 = 1 if the vehicle has 6 to 7 cylinders; **(6 to 7)**
  cyl_8p = 1 if the vehicle has 8 or more than 10 cylinders. **(8 or more than 10)**

  The group of vehicles with 8 or more than 10 cylinders is used as the reference category. Coefficients estimated as negative and significant for the other groups of vehicle/number of cylinders will thus indicate that these groups run lower risks than those in the reference group.

- *Vehicle's type of fuel:* 3 dichotomous variables have been created:
  diesel = 1 if the vehicle uses diesel as fuel; **(Diesel)**
  essence = 1 if the vehicle uses gas as fuel; **(Gas)**
  carb_aut = 1 if the vehicle uses another type of fuel. **(Other)**

  The group of vehicles using diesel as fuel is considered the reference category. Coefficients estimated as negative and significant for the other groups of vehicle/fuel will thus indicate that these groups of vehicles run lower risks than those in the reference group.

- *Maximum number of axles:* 7 dichotomous variables have been created:
  ess_0 = 1 if the maximum number of axles does not apply to this type of vehicle;
  ess_2 = 1 if the vehicle has two axles and a mass of between 3,000 and 4,000 kg; **(2 axles (3,000 to 4,000 kg))**

ess_2p = 1 if the vehicle has two axles and a mass higher than 4,000 kg; **(2 axles (4,000 kg and more))**
ess_3 = 1 if the vehicle has a maximum of three axles; **(3 axles)**
ess_4 = 1 if the vehicle has a maximum of four axles; **(4 axles)**
ess_5 = 1 if the vehicle has a maximum of five axles; **(5 axles)**
ess_6p = 1 if the vehicle has six or more axles. **(6 axles or more)**

The group of vehicles with two axles and a mass of between 3,000 and 4,000 kg is used as the reference category. Coefficients estimated as positive and significant for the other groups of vehicles/maximum-axle support will thus indicate that these groups of vehicles run lower risks that those in the reference group.

- *Vehicle's type of use:* 3 dichotomous variables for vehicles transporting goods have been created:
  compr = 1 if the vehicle is meant for commercial use, including the transportation of goods without a CTQ permit; **(Commercial use)**
  tbrgn = 1 if the vehicle is meant for the transportation of non-bulk goods that require a CTQ permit; **(Other than bulk goods)**
  tbrvr = 1 if the vehicle is meant for the transportation of bulk goods. **(Bulk goods)**

  The group of vehicles transporting bulk goods is used as the reference category. Coefficients estimated as negative and significant for other groups of vehicles/types-of-use will thus indicate that these groups of vehicles run lower risks than the reference group.

- Five variables have been created to measure the number of convictions per vehicle accumulated in the year preceding year *t* by one or more drivers:
  - Number of violations for speeding per vehicle committed the year preceding year *t*. A positive sign is predicted because more speeding violations should, on average, generate more accidents. **(For speeding)**
  - Number of violations for driving with a suspended license per vehicle committed the year preceding year *t*. A positive sign is predicted because more driving with a suspended license should, on average, generate more accidents. **(Suspended license)**
  - Number of violations for running a red light per vehicle committed the year preceding year *t*. A positive sign is predicted because more incidences of running a red light should, on average, generate more accidents. **(For running a red light)**
  - Number of violations for failure to respect a stop sign or a signal from a traffic officer per vehicle committed the year preceding year *t*. A positive sign is predicted because more incidents of failure to respect a stop sign or a signal from a traffic officer should, on average, generate more accidents. **(For ignoring a stop sign)**
  - Number of violations for failure to wear a seat belt per vehicle committed the year preceding year *t*. A positive sign is predicted because more incidents of failing to wear a seat belt should, on average, generate more accidents. **(Not wearing a seat belt)**

Appendix 3 – Estimation of the Hausman model

Table 7b: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991-1998 period (fleet of two trucks or more and trucks with two periods or more), with the Hausman, the gamma-Dirichlet and the Hierarchical I models.

Explanatory variables	Hausman model		Gamma-Dirichlet model		Hierarchical I	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
Constant	-0.1254	0.0819	-3.9070*	0.0573	-3.6435*	0.0548
Number of years as a fleet	-0.0436*	0.0031	-0.0464*	0.0044	-0.0260*	0.0040
Sector of activity in 1998						
Other sector	-0.2484*	0.0929	-0.1426	0.1163	-0.2681	0.1118
General public trucking	0.1003*	0.0252	0.1685*	0.0304	0.0727	0.0324
Bulk public trucking	Reference group		Reference group		Reference group	
Private trucking	0.1574*	0.0213	0.2290*	0.0256	0.2007*	0.0267
Short-term rental firm	0.4480*	0.0336	0.5633*	0.0483	0.1391*	0.0450
Size of fleet						
2	Reference group		Reference group		Reference group	
3	0.1260*	0.0180	0.0801*	0.0205	0.1287*	0.0194
4 to 5	0.1941*	0.0172	0.1385*	0.0205	0.2104*	0.0197
6 to 9	0.2798*	0.0171	0.2137*	0.0210	0.3114*	0.0209
10 to 20	0.3617*	0.0166	0.2937*	0.0209	0.4510*	0.0221
21 to 50	0.3574*	0.0177	0.3010*	0.0223	0.6678*	0.0258
More than 50	0.3591*	0.0167	0.3077*	0.0217	1.5852*	0.0261
Days in previous year	1.6878*	0.0300	2.0537*	0.0300	1.7784*	0.0298
Violations						
For overload	0.1216*	0.0117	0.0966*	0.0115	0.0809*	0.0119
For excessive size	0.1456***	0.0883	0.1480***	0.0860	0.1448	0.0884
For poorly secured cargo	0.2522*	0.0363	0.2054*	0.0354	0.1826*	0.0365
Not respect service hours	0.2585*	0.0663	0.1984*	0.0664	0.1984*	0.0678
No mechanical inspection	0.2383*	0.0308	0.1778*	0.0298	0.1575*	0.0307
For other reasons	0.2678*	0.0779	0.1754**	0.0743	0.2113*	0.0771
Type of vehicle use						
Commercial use	-0.1407*	0.0213	-0.1938*	0.0212	-0.1443*	0.0231
Other than bulk goods	-0.0513**	0.0244	-0.1148*	0.0243	-0.1159*	0.0268
Bulk goods	Reference group		Reference group		Reference group	
Type of fuel						
Diesel	Reference group		Reference group		Reference group	
Gas	-0.4089*	0.0145	-0.3973*	0.0136	-0.3441*	0.0153
Other	-0.3109*	0.0775	-0.3079*	0.0736	-0.4090*	0.0824
Number of cylinders						
1 to 5	0.3591*	0.0440	0.2167*	0.0403	0.3656*	0.0462
6 to 7	0.3778*	0.0136	0.3780*	0.0126	0.3540*	0.0143
8 or more than 10	Reference group		Reference group		Reference group	
Number of axles						
2 axles (3,000 to 4,000 kg)	-0.1620*	0.0210	-0.2916*	0.0208	-0.2898*	0.0233
2 axles (4,000 kg and more)	-0.1715*	0.0150	-0.2850*	0.0150	-0.2856*	0.0173

Explanatory variables	Hausman model		Gamma-Dirichlet model		Hierarchical I	
	Coefficient	Standard error	Coefficient	Standard error	Coefficient	Standard error
3 axles	-0.1559*	0.0151	-0.1278*	0.0149	-0.1641*	0.0170
4 axles	-0.1896*	0.0199	-0.1321*	0.0190	-0.1590*	0.0222
5 axles	-0.2182*	0.0173	-0.1973*	0.0174	-0.1914*	0.0194
6 axles or more	Reference group		Reference group		Reference group	
Number of violations						
For speeding	0.2585*	0.0105	0.1946*	0.0103	0.1849*	0.0107
Suspended license	0.4494*	0.0426	0.3830*	0.0422	0.3740*	0.0430
For running a red light	0.3838*	0.0247	0.3094*	0.0239	0.2815*	0.0246
For ignoring a stop sign	0.4264*	0.0267	0.3597*	0.0258	0.3150*	0.0266
Not wearing a seat belt	0.2044*	0.0304	0.1568*	0.0294	0.1362*	0.0303
Observation period						
1991	0.0187	0.0251	0.0760**	0.0332	0.1990*	0.0308
1992	-0.0183	0.0226	0.0548***	0.0293	0.1085*	0.0272
1993	-0.0837*	0.0208	0.0806*	0.0259	0.0223	0.0244
1994	-0.0201	0.0190	0.1845*	0.0226	0.0569*	0.0215
1995	0.0014	0.0175	0.2073*	0.0197	0.0581*	0.0190
1996	-0.0426*	0.0165	0.1198*	0.0175	-0.0201	0.0171
1997	-0.1583*	0.0163	-0.0791*	0.0163	-0.1613*	0.0163
1998	Reference group		Reference group		Reference group	
$\hat{a}$	56.9383*	3.4587				
$\hat{b}$	1.8274*	0.0384				
$\hat{\delta}$					0.8168*	0.0250
$\widehat{\beta c}_0$					3.0504*	0.0912
$\hat{\delta}^*$					4.7817*	1.4580
$\hat{\beta}^*$					10.1139*	3.1455
$\hat{\nu}$			2.0086*	0.0422		
$\hat{\kappa}$			12.6597*	0.2508		
$\hat{\rho}$			4.6690*	0.3102		
Number of observations	456,177		456,177		456, 177	
Number of trucks	111,106		111,106		111,106	
Log likelihood	-197,165		-197,116		-193,026	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$					0.7531	

* Significant at 1%; ** Significant at 5%; *** Significant at 10%

Online appendix

Gamma/Dirichlet model

Most of the econometric models applied to discrete (or count) variables start from the Poisson distribution, where the probability of truck *i* of fleet *f* being involved in y_{fit} accidents in period *t* can be represented by the following expression

$$P\left(Y_{fit} \mid \lambda_{fit}\right) = \frac{e^{-\lambda_{fit}}\left(\lambda_{fit}\right)^{y_{fit}}}{\Gamma\left(y_{fit}+1\right)}.$$

To simultaneously take into account both the firm effect and the time effect, suppose that $\lambda_{fit} = \gamma_{fit}\left(\alpha_f \theta_{(f)i} \eta_{(fi)t}\right)$ with $\gamma_{fit} = e^{X_{fit}\beta}$. The vector $X_{fit} = \left(x_{fit1}, \cdots, x_{fitp}\right)$ represents the *p* characteristics of truck *i* of fleet *f* observed in period *t*. This vector contains specific information about the vehicle and other specific information about the fleet. β is a vector of *p* parameters to be estimated. Let α_f be the random effect associated with fleet *f* (i.e. the risk or non-observable characteristics attributable to the fleet), whereas $\theta_{(f)i}$ is the random effect of truck *i* of fleet *f* where $\sum_{i=1}^{I_f} \theta_{(f)i} = 1$ where I_f is the number of vehicles in fleet *f*. Finally, $\eta_{(fi)t}$ is the random effect of period *t* of truck *i* of fleet *f* such as $\sum_{t=1}^{T_i} \eta_{(fi)t} = 1$ where T_i is the number of periods for truck *i*.

Angers et al (2018) make the three following hypotheses. The parameter α_f follows a gamma distribution of parameters $\left(\sum_{i=1}^{I_f} T_i\,\kappa^{-1}, \kappa^{-1}\right)$. The vector $\underset{\sim}{\theta}_{(f)} = \left(\theta_{(f)1}, \theta_{(f)2}, \cdots, \theta_{(f)I_f}\right)$ follows a Dirichlet distribution of parameters $\left(\nu_{(f)1}, \nu_{(f)2}, \cdots, \nu_{(f)I_f}\right)$ and the vector $\underset{\sim}{\eta}_{(fi)} = \left(\eta_{(fi)1}, \eta_{(fi)2}, \cdots, \eta_{(fi)T_i}\right)$ follows a Dirichlet distribution of parameters $\left(\rho_{(fi)1}, \rho_{(fi)2}, \cdots, \rho_{(fi)T_i}\right)$ where T_i is the number of periods of the vehicle *i*. Using these assumptions, the following expression for accident distribution is obtained:

$$P\left(Y_{f11},\cdots,Y_{fI_fT_{I_f}} \mid \underset{\sim}{\eta}_{(fi)}\right)$$

$$= \frac{\left[\Gamma\left(S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)\right]\left[\left(\kappa^{-1}\right)^{\sum_{i=1}^{I_f}T_i\kappa^{-1}}\right]\left[\Gamma\left(\sum_{i=1}^{I_f}\nu_{(f)i}\right)\right]\left[\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\left(\gamma_{fit}\eta_{(fi)t}\right)^{y_{fit}}\right]}{\left[\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\Gamma\left(y_{fit}+1\right)\right]\left[\Gamma\left(\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)\right]\left[\prod_{i=1}^{I_f}\Gamma\left(\nu_{(f)i}\right)\right]}\int\cdots\int\frac{\prod_{i=1}^{I_f}\left(\theta_{(f)i}\right)^{S_i+\nu_{(f)i}-1}}{\left(\kappa^{-1}+\sum_{i=1}^{I_f}\theta_{(f)i}\sum_{t=1}^{T_i}\gamma_{fit}\eta_{(fi)t}\right)^{S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}}}\,d\underset{\sim}{\theta}_{(f)}.$$

(A2)

By supposing that $\gamma_{fit}=\overline{\gamma}_{fi}\forall t=1,\cdots,T_i$ where $\overline{\gamma}_{fi}=\frac{1}{T_i}\sum_{i=1}^{T_i}\gamma_{fit}$, the integral of equation (A2) can be approximated. Separating the vehicles into two groups and defining $G_1=1,\cdots,g_1$ as the set of all vehicles of the first group with $\overline{\gamma}_{g_1}=\frac{\sum_{i=1}^{g_1}\overline{\gamma}_{fi}}{g_1}$, and $G_2=g_1+1,\cdots,I_f$, as the set of all vehicles of the second group with $\overline{\gamma}_{g_2}=\frac{\sum_{i=g_1+1}^{I_f}\overline{\gamma}_{fi}}{g_2}$, the integral of equation (A2) thus becomes:

$$\int\cdots\int\frac{\left[\prod_{i=1}^{g_1}\left(\theta_{(f)i}\right)^{S_i+\nu_{(f)i}-1}\right]\left[\prod_{i=g_1+1}^{I_f}\left(\theta_{(f)i}\right)^{S_i+\nu_{(f)i}-1}\right]}{\left(\kappa^{-1}+\overline{\gamma}_{g_1}\sum_{i=1}^{g_1}\theta_{(f)i}+\overline{\gamma}_{g_2}\sum_{i=g_1+1}^{I_f}\theta_{(f)i}\right)^{S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}}}\,d\underset{\sim}{\theta}_{(f)}. \qquad \text{(A3)}$$

By integrating one obtains:

$$=\frac{\prod_{i=1}^{I_f}\Gamma\left(S_i+\nu_{(f)i}\right)}{\left[\Gamma\left(\sum_{i=1}^{I_f}S_i+\nu_{(f)i}\right)\right]\left[\left(\kappa^{-1}+\overline{\gamma}_{g_2}\right)^{S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}}\right]}\,{}_2F_1\left(\sum_{i=1}^{g_1}S_i+\nu_{(f)i},S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1},\sum_{i=1}^{I_f}S_i+\nu_{(f)i},\left(\frac{\overline{\gamma}_{g_2}-\overline{\gamma}_{g_1}}{\kappa^{-1}+\overline{\gamma}_{g_2}}\right)\right).$$

(A4)

Thus, by replacing the integral in equation (A3) with its value given in (A4) the following approximation for $P\left(Y_{f11},\cdots,Y_{fI_fT_{I_f}} \mid \underset{\sim}{\eta}_{(fi)}\right)$ in (A2) is equal to:

$$\frac{\Gamma\left(S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)\left(\kappa^{-1}\right)^{\sum_{i=1}^{I_f}T_i\kappa^{-1}}\Gamma\left(\sum_{i=1}^{I_f}\nu_{(f)i}\right)\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\left(\gamma_{fit}\eta_{(fi)t}\right)^{y_{fit}}}{\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\Gamma\left(y_{fit}+1\right)\Gamma\left(\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)\prod_{i=1}^{I_f}\Gamma\left(\nu_{(f)i}\right)}\left(\frac{\prod_{i=1}^{I_f}\Gamma\left(S_i+\nu_{(f)i}\right)}{\Gamma\left(\sum_{i=1}^{I_f}S_i+\nu_{(f)i}\right)}\right)\frac{1}{\left(\kappa^{-1}+\overline{\gamma}_{g_2}\right)^{S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}}}$$

$$\times {}_2F_1\left(\sum_{i=1}^{g_1}S_i+\nu_{(f)i},S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1},\sum_{i=1}^{I_f}S_i+\nu_{(f)i},\left(\frac{\overline{\gamma}_{g_2}-\overline{\gamma}_{g_1}}{\kappa^{-1}+\overline{\gamma}_{g_2}}\right)\right) \tag{A5}$$

where ${}_2F_1$ is a hypergeometric function whose value is equal to:

$$1+\sum_{\ell=1}^{\infty}\left[\frac{\left(\sum_{i=1}^{g_1}S_i+\nu_{(f)i}\right)^{[\ell]}\left(S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)^{[\ell]}}{\left(\sum_{i=1}^{I_f}S_i+\nu_{(f)i}\right)^{[\ell]}}\frac{\left(\frac{\overline{\gamma}_{g_2}-\overline{\gamma}_{g_1}}{\kappa^{-1}+\overline{\gamma}_{g_2}}\right)^{\ell}}{\ell!}\right],$$

with $h^{[\ell]}=h\left(h+1\right)\cdots\left(h+\ell+1\right)$, an increasing factorial function. More substitutions yield:

$$\left[\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\frac{\left(\gamma_{fit}\right)^{y_{fit}}}{\Gamma\left(y_{fit}+1\right)}\right]\left[\frac{\Gamma\left(S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)}{\Gamma\left(\sum_{i=1}^{I_f}T_i\kappa^{-1}\right)}\right]\left[\frac{\left(\kappa^{-1}\right)^{\sum_{i=1}^{I_f}T_i\kappa^{-1}}}{\left(\kappa^{-1}+\overline{\gamma}_{g_2}\right)^{S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1}}}\right]\left[\frac{\Gamma\left(\sum_{i=1}^{I_f}\nu_{(f)i}\right)}{\Gamma\left(\sum_{i=1}^{I_f}S_i+\nu_{(f)i}\right)}\right]\left[\frac{\prod_{i=1}^{I_f}\Gamma\left(S_i+\nu_{(f)i}\right)}{\prod_{i=1}^{I_f}\Gamma\left(\nu_{(f)i}\right)}\right]$$

$$\times\left[\frac{\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\Gamma\left(y_{fit}+\rho_{(fi)t}\right)}{\prod_{i=1}^{I_f}\prod_{t=1}^{T_i}\Gamma\left(\rho_{(fi)t}\right)}\right]\left[\frac{\prod_{i=1}^{I_f}\Gamma\left(\sum_{t=1}^{T_i}\rho_{(fi)t}\right)}{\prod_{i=1}^{I_f}\Gamma\left(S_i+\sum_{t=1}^{T_i}\rho_{(fi)t}\right)}\right]\times {}_2F_1\left(\sum_{i=1}^{g_1}\left(S_i+\nu_{(f)i}\right),S_0+\sum_{i=1}^{I_f}T_i\kappa^{-1},\sum_{i=1}^{I_f}\left(S_i+\nu_{(f)i}\right),\left(\frac{\overline{\gamma}_{g_2}-\overline{\gamma}_{g_1}}{\kappa^{-1}+\overline{\gamma}_{g_2}}\right)\right). \tag{A6}$$

Letting $\nu_{(f)i}=\nu\ \forall i$ and $\rho_{(fi)t}=\rho\ \forall t$, we then used the maximum likelihood method to estimate the unknown parameters, ν,κ^{-1},ρ and β of the log likelihood function of the model.

Other estimations and results

Different values of *K*

Table O1 presents the results of the Hierarchical I random model with *K*=5 in columns 2 and 3 and with *K*=10 in columns 4 and 5. Table O2 gives the results for *K*=11 and 12. The log likelihood is -195,069 at *K*=5 and increases to -193,026 at *K*=13, a value very close to the log likelihood values at *K*=10 (193,649) and at *K*=12 (-193,226).

From *K*=10 to *K*=13 the estimate of the constant varies from -3.6260 to -3.6435.

The number of years as a fleet on December 31st varies from-0.0263 to -0.0260.

Sector of activity:

- Other varies from -0.2870 to -0.2681;
- General public trucking varies from 0.0591 to 0.0727;
- Private trucking varies from 0.1802 to 0.2007;
- Short-term rental firm varies from 0.1342 to 0.1391.

Size of fleet:

- 3 trucks varies from 0.1285 to 0.1287;
- 4 – 5 trucks varies from 0.2084 to 0.2104;
- 6 – 9 trucks varies from 0.3039 to 0.3114;
- 10 – 20 trucks varies from 0.4222 to 0.4510;
- 21 – 50 trucks varies from 0.5721 to 0.6678;
- More than 50 trucks varies from 1.3443 to 1.5852.

As the size of the fleet increases the estimate difference increases.

It is interesting to note that the coefficient estimates corresponding to the number of violations of trucking standards, number of violations with demerit points, observation period, and truck characteristics are fairly stable from *K*=10 to *K*=13.

$\hat{\delta}$ varies from 0.6685 to 0.8168

$\widehat{\beta c_0}$ varies from 2.4790 to 3.0504

$\hat{\delta}^*$ varies from 5.4334 to 4.7817

$\hat{\beta}^*$ varies from 10.1362 to 10.1139

Table O1: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991-1998 period (fleet of two trucks or more and trucks with two periods or more), and hierarchical random effect models with K = 5 and K = 10.

Explanatory variables	Hierarchical I/K=5		Hierarchical I/K=10	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.5814*	0.0509	-3.6260*	0.0542
Number of years as a fleet on December 31st	-0.0304*	0.0036	-0.0263*	0.0039
Sector of activity in 1998				
Other sector	-0.2988*	0.0988	-0.2870*	0.1055
General public trucking	0.0774*	0.0287	0.0591	0.0302
Bulk public trucking	Reference group		Reference group	
Private trucking	0.1435*	0.0239	0.1802*	0.0267
Short-term rental firm	0.2420*	0.0417	0.1342*	0.0439
Size of fleet				
2	Reference group		Reference group	
3	0.1218*	0.0188	0.1285*	0.0193
4 to 5	0.1877*	0.0186	0.2084*	0.0207
6 to 9	0.2627*	0.0193	0.3039*	0.0195
10 to 20	0.3303*	0.0194	0.4222*	0.0217
21 to 50	0.3565*	0.0218	0.5721*	0.0247
More than 50	0.8368*	0.0207	1.3443*	0.0239
Number of days authorized to drive in the previous year	1.7605*	0.0298	1.7792*	0.0298
Number of violations of trucking standards in the previous year				
For overload	0.1006*	0.0121	0.0841*	0.0120
For excessive size	0.1512	0.0898	0.1466	0.0889
For poorly secured cargo	0.2131*	0.0377	0.1865*	0.0368
For failure to respect service hours	0.2141*	0.0699	0.1988*	0.0682
For failure to pass mechanical inspection	0.1848*	0.0314	0.1614*	0.0309
For other reasons	0.2188*	0.0783	0.2135*	0.0776
Type of vehicle use				
Commercial use including transport of goods without C.T.Q. permit	-0.1645*	0.0224	-0.1503*	0.0232
Transport of other than "bulk" goods	-0.0901*	0.0259	-0.1147*	0.0268
Transport of "bulk" goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.3946*	0.0150	-0.3502*	0.0152
Other	-0.3795*	0.0804	-0.4084*	0.0817
Number of cylinders				
1 to 5 cylinders	0.3989*	0.0460	0.3694*	0.0463
6 to 7 cylinders	0.3755*	0.0141	0.3604*	0.0143
8 or more than 10 cylinders	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.2790*	0.0225	-0.2853*	0.0231
2 axles (more than 4,000 kg)	-0.2453*	0.0165	-0.2808*	0.0172
3 axles	-0.1929*	0.0166	-0.1741*	0.0169
4 axles	-0.2265*	0.0185	-0.1730*	0.0219

Explanatory variables	Hierarchical I/*K*=5		Hierarchical I/*K*=10	
	Coefficient	Standard error	Coefficient	Standard error
5 axles	-0.2206*	0.0185	-0.1982*	0.0193
6 axles or more	Reference group		Reference group	
Number of violations with demerit points in the previous year				
For speeding	0.2174*	0.0110	0.1893*	0.0108
For driving with a suspended license	0.4035*	0.0438	0.3775*	0.0433
For running a red light	0.3231*	0.0252	0.2871*	0.028
For ignoring a stop sign or traffic officer	0.3487*	0.0272	0.3199*	0.0268
For not wearing a seat belt	0.1645*	0.0307	0.1403*	0.0304
Observation period				
1991	0.1440*	0.0282	0.1913*	0.0301
1992	0.0752*	0.0249	0.1053*	0.0266
1993	-0.0073	0.0226	0.0200	0.0239
1994	0.0342	0.0203	0.0550*	0.0212
1995	0.0429	0.0183	0.0574*	0.0188
1996	-0.0228	0.0168	-0.0184	0.0170
1997	-0.1540*	0.0163	-0.1576*	0.0163
1998	Reference group		Reference group	
$\hat{\delta}$	0.3668*	0.0172	0.6685*	0.0240
$\widehat{\beta c}_0$	1.3059*	0.0566	2.4790*	0.0835
$\hat{\delta}^*$	4.5711*	0.6038	5.4334*	1.4223
$\hat{\beta}^*$	9.2089*	1.2541	10.1362*	2.7183
Number of observations	456, 177		456, 177	
Log likelihood	-195,069		-193,649	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.5663		0.7126	

* Significant at 1%.

Table O2: Estimation of the parameters of the distribution of the number of annual truck accidents for the period of 1991-1998 (fleet of two trucks or more and trucks with two periods or more), and hierarchical random effect models with K = 11 and K = 12.

Explanatory variables	Hierarchical I/K=11		Hierarchical I/K=12	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.6335*	0.0544	-3.6393*	0.0546
Number of years as a fleet on December 31st	-0.0262*	0.0039	-0.0261*	0.0040
Sector of activity in 1998				
Other sector	-0.2782*	0.1074	-0.2718	0.1095
General public trucking	0.0643	0.0322	0.0687	0.0323
Bulk public trucking	Reference group		Reference group	
Private trucking	0.1885*	0.0268	0.1954*	0.0269
Short-term rental firm	0.1372*	0.0440	0.1388*	0.0443
Size of fleet				
2	Reference group		Reference group	
3	0.1285*	0.0194	0.1286*	0.0194
4 to 5	0.2091*	0.0208	0.2097*	0.0196
6 to 9	0.3065*	0.0219	0.3081*	0.0209
10 to 20	0.4321*	0.0219	0.4316*	0.0220
21 to 50	0.6061*	0.0250	0.6380*	0.0254
More than 50	1.4277*	0.0245	1.5085*	0.0253
Number of days authorized to drive in the previous year	1.7791*	0.0298	1.7787*	0.0298
Number of violations of trucking standards in the previous year				
For overload	0.0826*	0.0120	0.0816*	0.0119
For excessive size	0.1456	0.0887	0.1451	0.0885
For poorly secured cargo	0.1847*	0.0367	0.1835*	0.0366
For failure to respect service hours	0.1984*	0.0680	0.1983*	0.0679
For failure to pass mechanical inspection	0.1597*	0.0308	0.1584*	0.0308
For other reasons	0.2125*	0.0774	0.2118*	0.0773
Type of vehicle use				
Commercial use including transport of goods without C.T.Q. permit	-0.1478*	0.0232	-0.1458*	0.0231
Transport of other than "bulk" goods	-0.1149*	0.0268	-0.1153*	0.0268
Transport of "bulk" goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.3479*	0.0153	-0.3459*	0.0153
Other	-0.4088*	0.0819	-0.4090*	0.0821
Number of cylinders				
1 to 5 cylinders	0.3684*	0.0463	0.3672*	0.0462
6 to 7 cylinders	0.3580*	0.0143	0.3559*	0.0143
8 or more than 10 cylinders	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.2872*	0.0232	-0.2887*	0.0232
2 axles (more than 4,000 kg)	-0.2828*	0.0172	-0.2843*	0.0173
3 axles	-0.1700*	0.0170	-0.1667*	0.0170
4 axles	-0.1676*	0.0220	-0.1629*	0.0221

Explanatory variables	Hierarchical I/K=11		Hierarchical I/K=12	
	Coefficient	Standard error	Coefficient	Standard error
5 axles	-0.1956*	0.0193	-0.1933*	0.0194
6 axles or more	Reference group		Reference group	
Number of violations with demerit points the previous year				
For speeding	0.1874*	0.0170	0.1860*	0.0170
For driving with a suspended license	0.3761*	0.0432	0.3749*	0.0431
For running a red light	0.2847*	0.0247	0.2829*	0.0247
For ignoring a stop sign or traffic officer	0.3178*	0.0267	0.3162*	0.0266
For not wearing a seat belt	0.1386*	0.0304	0.1373*	0.0304
Observation period				
1991	0.1941*	0.0303	0.1966*	0.0305
1992	0.1064*	0.0268	0.1075*	0.0270
1993	0.0212	0.0241	0.0223	0.0242
1994	0.0557*	0.0213	0.0563*	0.0214
1995	0.0577*	0.0189	0.0579*	0.0189
1996	-0.0190	0.0171	-0.0195	0.0171
1997	-0.1588*	0.0163	-0.1601*	0.0163
1998	Reference group		Reference group	
$\hat{\delta}$	0.7250*	0.0245	0.7748*	0.0248
$\widehat{\beta c}_0$	2.6635*	0.0853	2.8531*	0.0878
$\hat{\delta}^*$	5.2001*	1.4203	5.0178*	1.4611
$\hat{\beta}^*$	10.0707*	2.8101	10.1418*	3.0145
Number of observations:	456, 177		456, 177	
Log likelihood	-193,430		-193,226	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.7270		0.7405	

* Significant at 1%.

Fleet of five trucks

We estimated the models for fleets of 5 trucks with 1,257 fleets and 26,562 observations. As mentioned, all fleets must have 2 trucks or more per year, so those 1,257 fleets have 5 trucks, but they can have 2, 3, 4 or 5 trucks on the road per year. From those 1,257 fleets, only 6.4% of them had 5 trucks per year throughout the observation period.

Table O3 gives the estimation results of the gamma/Dirichlet and the hierarchical model for the fleets of 5 trucks. The log likelihood values are similar for the two models, namely -10,571 (gamma/Dirichlet) and -10,468 (hierarchical). The 3 random effects parameters, κ, ρ and ν of the gamma/Dirichlet are significantly different from 0 while, for the hierarchical model, only 2 parameters over 4 for the random effects are significantly different from 0 at 1% level of significance.

Overall, the standard errors are quite similar from one model to another. However, for the coefficients there is a little more variation depending on the explanatory variable. The year coefficients have the most differences between the two models.

Number of violations of trucking standards in the previous year: The estimate standard errors are stable between the 2 models.

Number of violations with demerit points the previous year: All estimates are significant at 1% for the two models, except for not wearing the seat belt which is significant at 5% for gamma/Dirichlet and hierarchical models and for not wearing a seat belt for the gamma/Dirichlet model.

Fleet of ten trucks

For the 234 fleets of size 10 in Table O4, the log-likelihood values are also close (-4,382 for the gamma/Dirichlet and -4,310 for the hierarchical). Note that the latter did not converge with K = 13 but with K = 12 and two of the 4 parameters for the random effects are not significant and are very high. The hierarchical model had more difficulty converging than the gamma/Dirichlet model.

In general, the models are similar with a few exceptions, such as private trucking. Likewise, the coefficients for the years differ from one model to another but they are not significant regardless of the model.

The same applies to the sector of activity, particularly other sector and private trucking. And for the type of vehicle use.

Number of violations of trucking standards in the previous year: The estimate standard errors are stable between the 2 models. The coefficients for overload are significant at 5% for all two models. For poorly secured cargo, the coefficient is significant at 1% for all three models.

Number of violations with demerit points in the previous year: Coefficients for speeding, for running a red light and for not wearing a seat belt are significant at 1% for the two models. The coefficient for ignoring a stop sign is significant at 5% for all models.

Fleet of 20 to 50 trucks

There are 548 fleets with between 20 and 50 trucks. The results are in Table O5. The log likelihood for the two models are: gamma/Dirichlet -33,709 and hierarchical -33,348 a difference of 361.

For the gamma/Dirichlet model the 3 parameters κ, ρ and ν are significantly different from 0. The same is observed for the hierarchical model where the 4 parameters are significantly different from 0.

The observation period is significant at 1% for the year 1994 to 1997 and at 5% for 1993 for the gamma/Dirichlet and only for the year 1997 for the hierarchical model.

The standard error of the number of years as a fleet are very different between gamma/Dirichlet and hierarchical. The coefficient is very different from these two models for the private trucking firms. They also differ in the type of vehicle use.

Number of violations of trucking standards in the previous year: The estimated standard errors are stable between the two models. The coefficient for overload is significant at 1% for the gamma/Dirichlet and at 5% for the Hausman model. For excessive size, it is significant at 5% for the gamma/Dirichlet and hierarchical models. For poorly secured cargo it is significant at 1% for all three models.

Table O3: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991-1998 period (fleet of five trucks with two periods or more), gamma/Dirichlet model and hierarchical random effect models with K = 13.

Explanatory variables	Gamma/Dirichlet		Hierarchical/K=13	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.3895	0.2177	-3.1191*	0.2052
Number of years as a fleet on December 31st	-0.0729	0.0160	-0.0533*	0.0145
Sector of activity in 1998				
Other sector	-0.2873*	0.3476	-0.2483	0.3744
General public trucking	0.0264*	0.1172	-0.0023	0.1179
Bulk public trucking	Reference group		Reference group	
Private trucking	0.2108*	0.0814	0.1748**	0.0823
Short-term rental firm	0.1615	0.3867	0.2292	0.3440
Number of days authorized to drive in the previous year	2.0544*	0.1277	1.7170*	0.1269
Number of violations of trucking standards in the previous year				
For overload	0.0799	0.0493	0.0787	0.0505
For excessive size	0.3032	0.3687	0.2605	0.3845
For poorly secured cargo	0.1935	0.1473	0.2269	0.1530
For failure to respect service hours	0.3477	0.5354	0.2356	0.5181
For failure to pass mechanical inspection	0.1060	0.1119	0.1351	0.1145
For other reasons	0.0120	0.3452	0.1442	0.3591
Type of vehicle use				
Commercial use including transport of goods without C.T.Q. permit	-0.3070*	0.0723	-0.2693*	0.0761
Transport of other than "bulk" goods	0.0705	0.0993	0.0836	0.1038
Transport of "bulk" goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.5338*	0.0582	-0.5284*	0.0608
Other	-0.6991	0.4262	-0.5947	0.4447
Number of cylinders				
1 to 5 cylinders	0.2479	0.1395	0.2196	0.1590
6 to 7 cylinders	0.3031*	0.0533	0.2928*	0.0565
8 or more than 10 cylinders	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.1929**	0.0886	-0.1800	0.0923
2 axles (more than 4,000 kg)	-0.2596*	0.0675	-0.2726*	0.0706

Explanatory variables	Gamma/Dirichlet		Hierarchical/K=13	
	Coefficient	Standard error	Coefficient	Standard error
3 axles	-0.3421*	0.0640	-0.3815*	0.0666
4 axles	-0.0355	0.0912	-0.0580	0.0973
5 axles	-0.2363*	0.0773	-0.2093*	0.0802
6 axles or more	Reference group		Reference group	
Number of violations with demerit points the previous year				
For speeding	0.2285*	0.0443	0.2493*	0.0465
For driving with a suspended license	0.7191*	0.1488	0.6840*	0.1569
For running a red light	0.4416*	0.1020	0.4350*	0.1057
For ignoring a stop sign or traffic officer	0.2848**	0.1115	0.3051*	0.1165
For not wearing a seat belt	0.2430**	0.1087	0.2586**	0.1123
Observation period				
1991	-0.0052	0.1249	0.0540	0.1159
1992	-0.0662	0.1114	-0.0345	0.1038
1993	-0.0330	0.1007	-0.0987	0.0951
1994	0.1364	0.0895	0.0138	0.0856
1995	0.1244	0.0813	-0.0134	0.0789
1996	0.1535**	0.0738	0.0331	0.0728
1997	-0.0827	0.0715	-0.1357	0.0717
1998	Reference group		Reference group	
$\hat{\delta}$			0.5437*	0.1273
$\widehat{\beta c}_0$			1.8724*	0.2488
$\hat{\delta}^*$			3.2690**	1.5877
$\hat{\beta}^*$			3.1479	1.6313
$\hat{\nu}$	2.0747*	0.0851		
$\hat{\kappa}$	1.4618*	0.3000		
$\hat{\rho}$	1.1172*	0.1339		
Number of observations	26,562		26,562	
Number of fleets	1,257		1,257	
Log likelihood	-10,571		-10,468	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1+\widehat{\beta c}_0}$			0.6519	

* Significant at 1%. ** Significant at 5%.

Table O4: Estimation of the parameters of the distribution of the number of annual truck accidents for the period of 1991-1998 (fleet of ten trucks with two periods or more), gamma/Dirichlet model and hierarchical random effect models with K = 12.

Explanatory variables	Gamma/Dirichlet		Hierarchical/K=12	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.2595*	0.4522	-3.0187*	0.3687
Number of years as a fleet on December 31st	-0.1275*	0.0370	-0.1022*	0.0283
Sector of activity in 1998				
Other sector	-0.1735	0.9940	-0.0780	0.8398
General public trucking	0.2960	0.2082	0.2781	0.1893
Bulk public trucking	Reference group		Reference group	
Private trucking	0.2378	0.1603	0.1565	0.1498
Short-term rental firm	0.8082**	0.4122	0.8919*	0.2953
Number of days authorized to drive in the previous year	2.0706*	0.2031	1.7315*	0.2012
Number of violations of trucking standards in the previous year				
For overload	0.1510**	0.0687	0.1698**	0.0715
For excessive size	0.1696	0.3771	0.1880	0.3685
For poorly secured cargo	0.4599*	0.1559	0.4797*	0.1704
For failure to respect service hours	0.2802	0.1763	0.2461	0.1771
For failure to pass mechanical inspection	0.1924	0.1663	0.2741	0.1715
For other reasons	0.7187**	0.3634	0.4932	0.3834
Type of vehicle use				
Commercial use including transport of goods without C.T.Q. permit	-0.0422	0.1320	-0.0104	0.1341
Transport of other than “bulk” goods	0.1589	0.1627	0.1761	0.1648
Transport of “bulk” goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.2628*	0.0874	-0.2484*	0.0909
Other	0.5869	0.4008	0.6178	0.4035
Number of cylinders				
1 to 5 cylinders	0.0865	0.3110	0.3030	0.3269
6 to 7 cylinders	0.3278*	0.0784	0.3016*	0.0830
8 or more than 10 cylinders	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.2098	0.1346	-0.1308	0.1375
2 axles (more than 4,000 kg)	-0.3685*	0.1013	-0.3515*	0.1038
3 axles	-0.1982**	0.0979	-0.2600*	0.1002
4 axles	-0.0235	0.1255	-0.0981	0.1318
5 axles	-0.0842	0.1033	-0.1113	0.1061
6 axles or more	Reference group		Reference group	
Number of violations with demerit points the previous year				
For speeding	0.1982*	0.0654	0.2530*	0.0671
For driving with a suspended license	0.3171	0.2395	0.2783	0.2580
For running a red light	0.5358*	0.1414	0.5596*	0.1446
For ignoring a stop sign or traffic officer	0.4008**	0.1567	0.3941**	0.1593

Explanatory variables	Gamma/Dirichlet		Hierarchical/K=12	
	Coefficient	Standard error	Coefficient	Standard error
For not wearing a seat belt	0.4586*	0.1690	0.4999*	0.1749
Observation period				
1991	-0.1923	0.2716	-0.0803	0.2127
1992	-0.1614	0.2387	-0.0869	0.1892
1993	-0.2459	0.2107	-0.2819	0.1712
1994	0.0258	0.1807	-0.0891	0.1502
1995	0.1084	0.1533	-0.0367	0.1329
1996	-0.0191	0.1328	-0.1415	0.1225
1997	-0.1040	0.1193	-0.1579	0.1162
1998	Reference group		Reference group	
$\hat{\delta}$			0.5020*	0.1869
$\widehat{\beta c}_0$			1.8987*	0.4393
$\hat{\delta}^*$			58.7983	563.5169
$\hat{\beta}^*$			45.5502	438.3277
$\hat{\nu}$	2.4800*	0.1536		
$\hat{\kappa}$	2.2877*	0.8498		
$\hat{\rho}$	1.1521*	0.1909		
Number of observations	10,114		10,114	
Number of fleets	234		234	
Log likelihood	-4,382		-4,310	
$\widehat{\beta c} = \dfrac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$			0.6550	

* Significant at 1%.   ** Significant at 5%.

Table O5: Estimation of the parameters of the distribution of the number of annual truck accidents for the period from 1991-1998 (fleet of 20 to 50 trucks with two periods or more), gamma/Dirichlet model and hierarchical random effect models with K = 13.

Explanatory variables	Gamma/Dirichlet		Hierarchical I/K=13	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.9136*	0.2556	-3.2584*	0.1919
Number of years as a fleet on December 31st	-0.0101	0.0223	-0.0153	0.2168
Sector of activity in 1998				
Other sector	-0.8131	0.4909	-0.1509	0.3695
General public trucking	0.0615	0.1048	0.0151	0.0958
Bulk public trucking	Reference group		Reference group	
Private trucking	0.1822	0.1003	0.0281	0.0914
Short-term rental firm	0.4724**	0.1935	0.4152*	0.1250
Number of days authorized to drive in the previous year	2.1493*	0.0733	1.7280*	0.0721
Number of violations of trucking standards in the previous year				
For overload	0.0702*	0.0264	0.0521	0.0277
For excessive size	0.4828**	0.2048	0.4835**	0.2214
For poorly secured cargo	0.2100*	0.0769	0.2075*	0.0794
For failure to respect service hours	0.1135	0.1370	0.1491	0.1430
For failure to pass mechanical inspection	0.0026	0.0832	-0.0096	0.0868
For other reasons	0.0574	0.1692	0.1546	0.1755
Type of vehicle use				
Commercial use including transport of goods without C.T.Q. permit	-0.1299**	0.0558	-0.0796	0.0604
Transport of other than “bulk” goods	-0.0506	0.0589	-0.0013	0.0639
Transport of “bulk” goods	Reference group		Reference group	
Type of fuel				
Diesel	Reference group		Reference group	
Gas	-0.3424*	0.0321	-0.2988*	0.0358
Other	-0.3354**	0.1633	-0.4037**	0.1838
Number of cylinders				
1 to 5 cylinders	0.3325*	0.0996	0.3724*	0.1175
6 to 7 cylinders	0.2795*	0.0296	0.2792*	0.0334
8 or more than 10 cylinders	Reference group		Reference group	
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3913*	0.0520	-0.3240*	0.0575
2 axles (more than 4,000 kg)	-0.3337*	0.0356	-0.3059*	0.0399
3 axles	-0.1957*	0.0357	-0.2416*	0.0399
4 axles	-0.0765	0.0423	-0.1567*	0.0474
5 axles	-0.1040**	0.0427	-0.1152**	0.0456
6 axles or more	Reference group		Reference group	
Number of violations with demerit points in the previous year				
For speeding	0.1689*	0.0237	0.1864*	0.0253
For driving with a suspended license	0.2419**	0.1088	0.2874**	0.1151
For running a red light	0.1094	0.0606	0.1067	0.0634
For ignoring a stop sign or traffic officer	0.2109*	0.0615	0.2307*	0.0651

Explanatory variables	Gamma/Dirichlet Coefficient	Gamma/Dirichlet Standard error	Hierarchical I/K=13 Coefficient	Hierarchical I/K=13 Standard error
For not wearing a seat belt	0.0335	0.0805	0.0593	0.0844
Observation period				
1991	0.2665	0.1591	0.1150	0.1114
1992	0.2439	0.1376	0.0779	0.0970
1993	0.2714**	0.1166	-0.0016	0.0836
1994	0.3808*	0.0957	0.0695	0.0702
1995	0.3604*	0.0757	0.0649	0.0580
1996	0.2496*	0.0575	0.0170	0.0475
1997	-0.0178	0.0436	-0.1303*	0.0407
1998	Reference group		Reference group	
$\hat{\delta}$			0.7128*	0.0831
$\widehat{\beta c}_0$			2.0541*	0.1829
$\hat{\delta}^*$			3.3180*	1.0244
$\hat{\beta}^*$			3.5199*	1.1422
$\hat{\nu}$	0.6818*	0.0453		
$\hat{\kappa}$	3.7528*	0.0676		
$\hat{\rho}$	1.5306*	0.1460		
Number of observations	68,475		68,475	
Number of fleets	548		548	
Log likelihood	-33,709		-33,348	
$\widehat{\beta c} = \dfrac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$			0.7053	

* Significant at 1%. ** Significant at 5%.

Additional *t*- paired tests

We also run a *t*-paired test of difference for the posterior expected number of accidents at period *T*+1 for size of the fleet and a *t*-paired test of difference for the accumulated numbers of truck accidents during the period. The results are presented in Table O6.

From Table O6, the means of the difference of the posterior expected number of accidents at period *T*+1 are not statistically different from 0 at 1% for fleets of size 2, 3, 4 and 5. For the other sizes of fleets the means of the difference are statistically different from 0 at 1%. Again the means of the difference are small, varying from 0.00120 to 0.00408 with one exception for fleets of more than 50 trucks where the mean of the difference is 0.03360.

Table O7 presents premiums for a diesel truck that transports non-“bulk” goods, has 8 cylinders or more and 6 axles or more. Its expected number of accidents at period *T*+1 is 0.2225 but falls to

0.2174 if it belongs to a fleet of size 2. We can see from Table O7 that its expected number of accidents at period *T*+1 does not vary very much according to the size of the fleet.

Its estimated number of accidents at period *T*+1 is 0.1595 if it does not have any accumulated accidents over the past but increases to 0.5490 if it had more than 3 accidents in the past.

Table O6: Paired *t*-test

	N	**A**	**B**	**D**	**std**	***t*-value**	***p*-value**
Fleet size							
2	13,776	0.1088	0.1091	-0.00027	0.0193	-1.66	0.0962
3	9,609	0.1206	0.1203	0.00027	0.0215	1.22	0.2241
4 - 5	14,397	0.1330	0.1326	0.00048	0.0236	2.45	0.0145
6 - 9	15,364	0.1512	0.1500	0.00120	0.0247	6.00	<0.0001
10 - 20	17,506	0.1732	0.1711	0.00212	0.0268	10.46	<0.0001
21 - 50	15,042	0.1859	0.1818	0.00408	0.0313	15.97	<0.0001
> 50	25,412	0.2073	0.1737	0.03360	0.0633	84.62	<0.0001
Cumulated number of accidents							
0	69,219	0.1008	0.1057	-0.0049	0.0239	-53.48	<0.0001
1	25,772	0.1976	0.1750	0.0226	0.0318	114.00	<0.0001
2	9,498	0.2995	0.2595	0.0401	0.0454	86.03	<0.0001
3	3,765	0.3944	0.3487	0.0457	0.0610	45.99	<0.0001
> 3	2,852	0.5583	0.4961	0.0622	0.1138	29.20	<0.0001

A: Mean of posterior expected number of accidents at period *T*+1 from estimations results of Hierarchical model I
B: Mean of posterior expected number of accidents at period *T*+1 from estimations results of Hierarchical model II
D: Mean of the difference between columns A and B

Table O7: Premiums for a diesel truck that transports non-“bulk” goods, has 8 cylinders or more and 6 axles or more, using: the optimal estimated Hierarchical model I in Table 7; the optimal estimated Hierarchical model II in Table 7, by size of fleet and cumulated number of accidents during the period.

	N	Hierarchical II model $E[X_{ijT+1} \mid X_{ikt}]$	Premium	Hierarchical I model $E[X_{ijT+1} \mid X_{ikt}]$	Premium
Size of fleet					
2	327	0.2005	$2,005	0.2174	$2,174
3	229	0.2268	$2,268	0.2388	$2,388
4 to 5	567	0.2247	$2,247	0.2366	$2,366
6 to 9	838	0.2492	$2,492	0.2539	$2,539
10 to 20	1,499	0.2394	$2,394	0.2424	$2,424
21 to 50	2,059	0.2236	$2,236	0.2233	$2,233
More than 50	3,740	0.2335	$2,335	0.2043	$2,043
	9,259	0.2316	$2,316	0.2225	$2,225

	N	Hierarchical II model $E[X_{ijT+1} \mid X_{ikt}]$	Premium	Hierarchical I model $E[X_{ijT+1} \mid X_{ikt}]$	Premium
Cumulated number of accidents over the period					
0	4,757	0.1478	$1,478	0.1595	$1,595
1	2,598	0.2483	$2,483	0.2276	$2,276
2	1,102	0.3540	$3,540	0.3108	$3,108
3	481	0.4471	$4,471	0.3977	$3,977
More than 3	521	0.5953	$5,953	0.5490	$5,490

Posterior expected number of accidents and premiums at period T+1 from estimations in Table 7.

Table O8 presents the premiums for different characteristics of the truck and sizes of the fleet using the optimal estimations presented in Table 7. The estimated numbers of accidents at period T+1 vary according to the characteristics of the truck and of the size of the fleet it belongs. If it belongs to a fleet of size 2, its expected number of accidents at period T+1 is 0.1091 and increases to 0.1711 if it belongs to a fleet of size 10 to 20 trucks.

Table O8: Examples of premiums for a truck depending of the size of fleet and truck characteristics

		Hierarchical I model		Hierarchical II model	
	N	$E[X_{ijT+1} \vert X_{ikt}]$	Premium ($)	$E[X_{ijT+1} \vert X_{ikt}]$	Premium ($)
Size of fleet					
2	13,776	0.1088	1,088	0.1091	1,091
3	9,609	0.1206	1,206	0.1203	1,203
4 to 5	14,397	0.1330	1,330	0.1326	1,326
6 to 9	15,364	0.1512	1,512	0.1500	1,500
10 to 20	17,506	0.1732	1,732	0.1711	1,711
21 to 50	15,042	0.1859	1,859	0.1818	1,818
More than 50	25,412	0.2073	2,073	0.1737	1,737
Type of vehicle use					
Commercial use including transport of goods without C.T.Q. permit	82,798	0.1500	1,500	0.1400	1,400
Transport of non-“bulk” goods	19,470	0.2052	2,052	0.1977	1,977
Transport of “bulk” goods	8,838	0.1787	1,787	0.1786	1,787
Type of fuel					
Diesel	87,546	0.1789	1,789	0.1703	1,703
Gas	22,999	0.0988	988	0.0893	893
Other	561	0.1006	1,006	0.0923	923
Number of cylinders					
1 to 5 cylinders	1,784	0.1561	1,561	0.1485	1,485
6 to 7 cylinders	71,159	0.1908	1,908	0.1810	1,810
8 or more than 10 cylinders	38,163	0.1084	1,084	0.1014	1,014
Number of axles					
2 axles (3,000 to 4,000 kg)	15,960	0.1160	1,160	0.1097	1,097
2 axles (more than 4,000 kg)	31,747	0.1415	1,415	0.1343	1,343
3 axles	21,856	0.1594	1,594	0.1467	1,467
4 axles	7,377	0.1740	1,740	0.1578	1,578
5 axles	10,666	0.1652	1,642	0.1572	1,572
6 axles or more	23,500	0.2179	2,179	0.2109	2,109

Posterior expected number of accidents and premiums at period T+1 from estimations in Table 7