Journal of Applied Econometrics 38, 2, 242-259, March 2023

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

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

17 October 2022

Summary

We propose a count-data model with hierarchical random effects for the posterior insurance ratemaking of vehicles belonging to a fleet, by allowing random effects for the fleet, the vehicles, and time. We derive a simple closed-form ratemaking formula based on a hierarchical random-effects specification. We estimate the corresponding econometric model, and compute insurance premiums according to the past experience of both the vehicle and the fleet. Our model can be used in other count-data applications with random individual and common effects on events involving many agents having activities with a principal in a hierarchical principal-agent environment, such as in education, health care management, finance, and business firms.

Keywords: Hierarchical model, vehicle insurance pricing, posterior ratemaking, asymmetric information, hierarchical random effects, hierarchical principal-agent.

JEL codes: C23, C25, C55, G22.

Conflict of interest

The authors report no conflicts of interest. They alone are responsible for the content and writing of the paper.

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 Canada's SSHRC, NSERC, the Canada Foundation for Innovation (CFI), and the SAAQ is also acknowledged, as well as the contributions of Claire Boisvert and Mohamed Jabir. We thank the discussants at the EGRIE conference (2021), the SIAM conference on Financial Mathematics and Engineering (2021), and the 6th Canadian Conference in Applied Statistics (2021), for their insightful comments.

[†] Denise Desjardins, HEC Montréal. denise.desjardins@hec.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, Concordia University. yang.lu@concordia.ca.

1. Introduction

Since the groundbreaking contribution of Hausman et al. (1984), count-data models have become popular in research areas such as labor economics (Kim and Marschke, 2005), public economics (Englin and Shonkwiler, 1995), health economics (Winkelmann, 2004), marketing (Böckenholt, 1998), and transportation safety (Dionne et al., 1997), to name only a few. These models began to be used for insurance pricing with Dionne and Vanasse (1989, 1992). The proposed Poisson regression model with gamma-distributed random effects has the advantage of allowing for a closed-form forecasting formula, and it has since become the basis for bonus-malus pricing under asymmetric information. Since then, extensions have flourished (Frangos and Vrontos, 2001; Purcaru and Denuit, 2003; Frees and Valdez, 2011). See Boucher and Guillen (2009) and Pinquet (2013) for reviews of random-effects insurance pricing models.

Most of the count-panel-data models in the literature involve random or fixed effects that are indexed by the individual and are often time-invariant. In this paper we consider individual, common, and time effects. There are many applications where panel data are available in groups or clusters. This is typically the case for fleet insurance. Other examples include analyzing absenteeism with matched employer-employee data (see Kim and Marschke, 2005; Abodd and Kramarz, 1999; Dionne and Dostie, 2007), comparing the relative impact of school and family (or teacher) on children's educational scores (Chamberlain, 2013; Freeman and Viarengo, 2014), analyzing health care failures and successes at the hospital and patient levels (Ludwig et al., 2010), as well as analyzing various county-level economic variables across different states (Connolly, 2016). Our econometric model can be applied to these data modeling environments, with random individual and common effects on events involving many agents working for an intermediary who must report to a principal under asymmetric information (Holmstrom, 1982; Holmstrom and

Milgrom, 1987; Sung, 2015; Hubert, 2020). (See Online appendix OA1 for an informal presentation of such principal-agent model with an intermediary.)

A negligent fleet manager (intermediary) may not spend enough money on truck maintenance and might ask employees to drive fast to achieve on-time deliveries. Conversely, the drivers (agents) may also exceed the speed limit without informing the manager. For the insurer (principal), knowing about all the accidents involving the vehicles belonging to a same fleet is essential in order to develop a fair and incentivized pricing scheme that accounts for the safety efforts made by each actor. For the regulator (another principal), this type of model makes it possible to compute the optimal fines for various infractions (driver, fleet owner), to improve social welfare as it relates to road safety.

While most count-panel models have assumed time-invariant random or fixed effects, Hausman et al. (1984) argue that, just as in linear panel data models, in which the error term is usually individual- and time-specific, such a multiple-level of heterogeneity for the random effect is also desirable in a count-data context. Indeed, in a Poisson count model with time-invariant random effects, the marginal variance-to-mean ratio is not individual specific. Moreover, the random-effects term controls for both the characteristics of the marginal distribution (such as overdispersion) and the serial correlation between counts. Consequently, the presence or lack of overdispersion may lead to spurious conclusions concerning the serial correlation (Lee et al., 2020).

Nevertheless, including multiple-level random effects in count-data models with panel data is not straightforward due to the ensuing computational difficulties. Let us consider the contributions of Zeger and Karim (1991) and Chib and Winkelmann (2001), which propose Poisson models with correlated lognormal random effects. The estimation and forecasting of these models typically rely on computationally intensive simulation techniques. These methods could be extremely costly in an application such as ours, where observations are triply indexed (common, individual, time).

To our knowledge, the first contribution of a nonlinear random-effects panel-data model with individual (here, driver), common (here, fleet), and time effects is from Angers et al. (2018). They extend the Hausman et al. (1984) model to add a firm effect to the individual and time effects of event distributions. Unfortunately, the Angers et al. (2018) model has some weaknesses. In particular, the model specification is not self-consistent when the number of years increases or when the set of vehicles changes over time. For example, the forecast of counts of one new period for a new vehicle requires the introduction of two new Dirichlet distributions of different dimensions than those for the parameters estimation (see the end of Online appendix OA3). This makes the model difficult to interpret and raises issues about its appropriateness for forecasting purposes. Moreover, the model does not allow for a closed-form formula for either the likelihood function or the forecasting formula, which require high-dimensional Monte Carlo simulations or numerical approximations.

The goal of this paper is to propose a new and tractable count model with panel data based on hierarchical random-effects specifications. Rather than following the aforementioned literature on multivariate lognormal-type random effects or extending the model of Angers et al. (2018), we take an entirely different approach. More precisely, our model draws inspiration from the recent time series Poisson-gamma conjugacy literature, to construct stationary-count or positively-valued processes with tractable properties (Pitt and Walker, 2015; Gouriéroux and Lu, 2019). We extend this technique and propose a hierarchical model for the triple-indexed random effects in which the upper-level latent variables are count-valued, and show that this discreteness greatly simplifies the estimation and forecasting processes in the resulting count-panel data model.

The model we propose is a parametric model, in the sense that the joint distribution of all the random effects belonging to the same fleet is fully specified. This is in line with the approach of Hausman et al. (1984) and Angers et al. (2018), as well as most of the aforementioned contributions on count-panel data. Alternatively, in the insurance literature, Norberg (1986), Desjardins et al. (2001), Pinquet (2013, 2020), and Fardilha et al. (2016) propose semiparametric approaches that specify the correlation structure between different counts, instead of their full joint distribution. Instead of getting a conditional linear expectation of the future counts, which is typically a highly nonlinear but positive function, these papers rely on linear regression models to forecast the future claims counts, i.e., the premium of each vehicle is a linear function of the number of past claims for that vehicle, as well as the number of past claims for the entire fleet. There are three major difficulties with this (second-order) moment-based approach: i) the discreteness of the response variables is not sufficiently taken into account; ii) the positivity of the regression coefficients, and hence the positivity of the insurance premium, is not guaranteed in the linear forecasting formulas; and iii) the linear premium does not sufficiently distinguish between the responsibilities of the driver and the fleet manager.[1]

Our model also differs in at least two respects from a competing hierarchical model for fleet insurance, proposed by Antonio et al. (2010). First, their model, which is similar in spirit to that of Zeger and Karim (1991), is not applicable to fleets that have only one vehicle at any given

[1] Let us for instance consider two otherwise equal fleets that differ only in terms of the distribution of their claims. Claims in Fleet A are concentrated on a single vehicle, suggesting that the driver of that vehicle is to blame, whereas claims in Fleet B are uniformly scattered across different vehicles, suggesting a potential management failure. A linear premium formula would lead to the same vehicle premium for both fleets, but a fairer pricing scheme should arguably penalize Fleet B more than Fleet A.

point in time while our model is estimated for fleets of one vehicle or more.[2] Second, their model does not allow a closed form for the likelihood function nor for the forecasting formula. Indeed, their posterior premium formula (see their Table 3) depends on unobservable random effects at the time, vehicle, and fleet levels. These formulas are not directly usable unless the posterior joint distribution of the random effects is recovered using the Bayes rule. This is a highly complicated task because the dimension of the joint distribution is the number of fleets/years.

The rest of the paper is organized as follows. Section 2 introduces the econometric model. Section 3 computes the theoretical likelihood function. Section 4 derives the posterior insurance pricing formula. Section 5 estimates the econometric model, using vehicle fleet data, and compares the results to previous contributions from the literature. Section 6 applies the new pricing formula to the data and presents out-of-sample tests. Section 7 concludes the paper. Additional information is presented in the Online appendix.

2. The theoretical model

Consider I fleets of vehicles, where each vehicle is doubly indexed, by their fleet ID $i=1,...,I$ and by 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 dozens or even hundreds. Finally, each vehicle can be observed during up to T periods. Thus, the claim counts are triply indexed, $X_{i,j,t}$, $i=1,...,I$, $j=1,...,s_i$, $t=1,...,T$, and we

[2] Indeed, as Antonio et al. (2010) mention, "with only one vehicle per fleet, the vehicle and fleet level coincide."

denote by $\lambda_{i,j,t}$ the associated a priori score, that is, the marginal expectation of $X_{i,j,t}$ given all observable covariates.

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 (pmf) is equal to $p(x) = \frac{\Gamma(x+\delta)}{x!\Gamma(\delta)} (1-p)^{\delta} p^{x}$. We denote by $\gamma(\delta, c)$ 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 $(X_{i,j,k})$ is as follows:

- At the highest hierarchical level, the $(N_i)_i$ are independent and identically distributed (*iid*) random effects following the $NB(\delta, \beta c)$ distribution, where $\beta c < 1$.
- At the second level, the $(Z_{i,j})$ are also random effects that are conditionally *iid* with the $P(\beta \eta_{i,j})$ distribution, where the $(\eta_{i,j})_j$ are themselves conditionally *iid,* 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 $(\theta_{i,j,t})_t$ conditionally *iid,* given $Z_{i,j}$, with gamma distribution $\gamma(\delta^* + \beta^* Z_{i,j}, c^*)$.
- Finally, given $(\theta_{i,j,t})$, the claim counts $X_{i,j,t}$ are independent and Poisson $P(\theta_{i,j,t}\lambda_{i,j,t})$ distributed.

To summarize, Figure 1 presents the model's chain rule. We now describe the state space of these random variables' levels. At the third level, we use $\theta_{i,j,t}$, which is continuously valued, so that $X_{i,j,t}$ has the standard Poisson random-effects 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 Proposition 1.

Figure 1 here

Proposition 1 – Properties of the hierarchical model

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*

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

Proof. Properties 1 and 2 are direct consequences of the Poisson-gamma conjugacy, which is reviewed in Online appendix OA2. As for Property 3, we have the following, from the (co)variance decomposition formula:

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

□

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:

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

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). In this special case, this implies that the correlation at the lowest level, $Corr\left[\theta_{i,j,t-1}, \theta_{i,j,t}\right]$, and at the intermediate level, $Corr\left[\eta_{i,j}, \eta_{i,j-1}\right]$, 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 $Corr[\theta_{i,j,t-1}, \theta_{i,j,t}]$ can attain any value 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 a fleet effect nor an individual effect.

3. Likelihood function

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

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

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

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

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

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

where the subscript *i* indicates that this quantity is fleet dependent, and the notation *X* indicates the fact that $M_i(X)$ depends on all the observable counts $\left(X_{i,j,t}\right)_{j,t}$, for all vehicles *j* and periods *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 compute

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

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

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

where the summation is with respect to *z*, that is, all the possible values of $Z_{i,j}$, and the term $\frac{\Gamma\left(\delta + N_i + z\right)}{\Gamma\left(\delta + N_i\right) z!} \frac{(\beta c)^z}{(1 + \beta c)^{\delta + N_i + z}}$ is the conditional pmf 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

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

Finally, equation (6) becomes

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

Using approximation (7) for $\prod_{j=1}^{s_i}M_{i,j}\left(X,N_i\right)$ and (8), we get an approximation for $M_i\left(X\right)$, which in turn leads to an approximation of the likelihood function in (5). Thus, the model's set of parameters is 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, $\lambda_{i,j,t}$. 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.

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 is; on the other hand, the larger K is, the more computational effort the method requires. Fortunately, our framework should involve a limited computational cost, which allows us to take quite large values of K and, hence, attain a high approximation quality. Indeed, the approximation of $M_i\left(X\right)$ requires us to consider the first $K+1$ smallest possible values of N_i. For each value n, we need to compute $M_{i,j}\left(X,n\right)$ in parallel for different j. As a result, computing the contribution of fleet i to the likelihood function requires a multiple of $s_i\left(K+1\right)^2$ operations, 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 the triplets (i, j, t).

We can compare this model to the gamma-Dirichlet model of Angers et al. (2018), which assumes that counts $X_{i,j,t}$ for fleet i, vehicle j, at time t are conditionally independent and Poisson distributed $P\left(\lambda_{i,j,t}\theta_{i,j,t}\right)$, where the random effect $\theta_{i,j,t}$ is further decomposed into

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

with the fleet effect $\alpha_{1,i}$ following a gamma distribution of parameters $(\sum_{j=1}^{s_i} T_j\kappa^{-1}, \kappa^{-1})$ where T_j is the number of periods of vehicle j, and s_i is the number of vehicles in fleet i. Both the fleet/vehicle effect $\left(\alpha_{2,i,j}\right)_j$ of dimension s_i and the fleet/vehicle/time effect $\left(\alpha_{3,i,j,t}\right)_t$ of dimension T follow Dirichlet distributions of parameters ν and ρ, respectively. Details of the model are presented in Online appendix OA3. The major restrictions of this gamma-Dirichlet approach are the following: i) It involves a Dirichlet distribution of dimension s_i, which becomes cumbersome when the fleet is large. ii) The resulting likelihood function does not have a closed-form expression except when the number of vehicles s_i is equal to 2. Although the authors work out some approximation techniques that are computationally less intensive than typical MCMC algorithms (Chib and Winkelman, 2001), approximation errors cannot be ignored and could be large, especially for large fleets. 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 the 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 might lead to arbitrage opportunities. v) The correlation between the random effects is not as flexible as in the new model presented in this article.

4. Forecasting formula

Because of the discrete latent random effects' 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 pmf, mixing variable N_i :

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

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 pmf 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):

$$p\left(\left(z_j\right)_j \mid \left(x_{i,k,t}\right)_{k,t}\right)$$
$$\propto \sum_{n=0}^{\infty} \frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!} \left(\beta c\right)^n \left(1-\beta c\right)^{\delta} \prod_{j=1}^{s_i} \left[\frac{\Gamma\left(\delta+n+z_j\right)}{\Gamma\left(\delta+n\right)z_j!} \frac{\left(\beta c\right)^{z_j}}{\left(1+\beta c\right)^{\delta+n+z_j}} \prod_{t=1}^{T} \frac{\Gamma(\delta^*+\beta^* z_j + x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)} \frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j + x_{i,j,t}}} \right] \tag{11}$$

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 observed claim counts for all vehicles *k* of 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,

$$\sum_{z_1=0}^{\infty} \cdots \sum_{z_{s_i}=0}^{\infty} \text{RHS of equation (11)}$$
$$= \sum_{n=0}^{\infty} \frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!} \left(\beta c\right)^n \left(1-\beta c\right)^{\delta} \prod_{j=1}^{s_i} \sum_{z_j=0}^{\infty} \left[\frac{\Gamma\left(\delta+n+z_j\right)}{\Gamma\left(\delta+n\right)z_j!} \frac{\left(\beta c\right)^{z_j}}{\left(1+\beta c\right)^{\delta+n+z_j}} \prod_{t=1}^{T} \frac{\Gamma(\delta^*+\beta^* z_j + x_{i,j,t})}{\Gamma(\delta^*+\beta^* z_j)} \frac{(c^*)^{x_{i,j,t}}}{(1+c^*\lambda_{i,j,t})^{\delta^*+\beta^* z_j + x_{i,j,t}}} \right]$$
$$= \sum_{n=0}^{\infty} \frac{\Gamma\left(\delta+n\right)}{\Gamma\left(\delta\right)n!} \left(\beta c\right)^n \left(1-\beta c\right)^{\delta} \prod_{j=1}^{s_i} M_{i,j}\left(X,n\right), \tag{12}$$

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 $\left(X_{i,j,t}\right)_{j,t}$, the second-level random effects $\left(Z_{i,j}\right)_j$ are still conditionally independent, given N_i, and that both $M_{i,j}\left(X,n\right)$ and the term between the square brackets in equation (11) can be computed in parallel for a different *j*. Consequently, we also deduce the marginal posterior distribution of each individual $Z_{i,j}$, given $\left(X_{i,j,t}\right)_{j,t}$:

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

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

Again, the computation of the above conditional distribution requires a 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 $T+1$:

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

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 pmf of $Z_{i,j} \mid \left(x_{i,k,t}\right)_{k,t}$ given by equation (13) through

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

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 accident data

We have access to the files of the Quebec motor vehicle department (*Société de l'assurance automobile du Québec*, henceforth referred to as the SAAQ) to create a database for the 1991–1998 period. The SAAQ is in charge of road safety regulations in the province and is the public insurer for personal injuries from traffic accidents. It also has information on all truck accidents involving a police report. Our starting point is the whole population of fleets registered in Quebec during the 1991–1997 period. To be registered, fleets must own at least one truck that is not used for emergencies. In this study, we have fleets of any size. The data on fleets contains information on violations (with convictions) committed by fleet owners 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 vehicles' mechanical inspection records and from the record of violations with convictions and demerit points (for speeding, etc.), as well as data on all accidents involving a police report. These include all traffic accidents causing bodily injury and all accidents causing material damage, reported by the police in Quebec. A description of the control variables can be found in Online appendix OA4.

5.1 Descriptive statistics

The data contains 62,171 fleets with a follow-up over at least two periods. As noted in Table OA5.1 of the Online appendix OA5, approximately 1% of the 62,171 fleets have more than 20 trucks.

On average, a truck has 4.13 follow-up periods, ranging from 3.88 to 4.30, while the average for a fleet is 4.95, ranging from 4.06 to 7.20 years. In Table OA5.2 of Online appendix OA5, we observe that about a quarter of fleets and 11% of trucks have eight years of follow-up, which confirms the panel aspect of the data. We also note that there are 13,059 fleets (45,308 trucks) for which we have only two consecutive years of follow-up.

Table OA5.3 of Online appendix OA5 shows the distribution of fleet sizes by year. In 1991, there are 31,793 fleets. This number increases over time, for a total of 307,792 fleet-years over the observation period. Among the 307,792 fleet-years, 70.91% have only one vehicle. In Table OA5.4 of Online appendix OA5, the average number of truck accidents per fleet is lowest during the year 1997. We also observe that the year 1995 has the highest recorded average rate of truck accidents per fleet.

In Table OA5.4, we see there are 66,193 trucks in 1991, for a total of 678,331 truck-years, with a mean annual truck accident rate of 13.72%. We observe that the annual truck accident rate is also lower in 1997.[3] We verified and found that a historical mistake was made when the data were transmitted by the insurer in 1999. It would have been too costly to make the corrections for all trucks and fleets, so we decided to keep the data as documented in 1999. As we will see, this will affect the out-of-sample analysis for the year 1997 but should not impact the model estimation.

Traffic violations committed by drivers and fleet owners are usually powerful forecasters of truck accidents in the following year. Indeed, we observe in Table OA5.6 of Online appendix OA5 that the year *t* accident rate is an increasing function of the previous year's violations committed by the drivers and fleet owners.

[3] See Table OA5.5 for per fleet size.

5.2 Estimation results

We estimate the models with all observations. For all models, we use the exponential linear specification for the marginal expectation

$$\lambda_{i,j,t} = E\left[X_{i,j,t} \middle| Y_{i,j,t}\right] = \exp\left(\vartheta' Y_{i,j,t}\right),$$

where vector ϑ contains the regression parameters, and $Y_{i,j,t}$ denotes the vector of observable covariates.

The estimation results of the hierarchical random-effects model are presented in Table 1. To decrease the convergence time, we computed the first and second derivatives of the likelihood function of the hierarchical model to obtain the gradient and the hessian. We truncated the infinite summation at two values, K =18 and K =19.[4] We observe that the log likelihood values are very similar, so we decided to not proceed with a higher value of K. In the next section, we show there is no statistical difference between the two models. Additionally, Table 1 shows that both the estimated standard errors and parameters are very stable between the two estimations. Other estimations with different values of K (8, 10, 13, 15) are presented in Table OA5.9 and Table OA5.10 of Online appendix OA5. When K increases, we observe that the standard error estimates of the explanatory variables remain fairly stable. This is not always the case for all coefficient estimates, particularly those of the random effects.

Several control variables presented in Online appendix OA4 measure observable heterogeneity. Some of these variables (type of fuel, number of cylinders, etc.) are characteristics of the vehicles, whereas others (sector, fleet size, etc.) relate to the fleet. We also include the number of violations of the trucking road-safety code by the fleet owner in the year before the accidents and the number of road-safety code violations leading to demerit points for the driver in

[4] Houston and Rossi (2017) used K=8, which is the largest truncation order we are aware of.

the year before the accidents**.** All coefficients of these variables, with three exceptions in Table 1, are significant at 1%. In fact, all coefficients that are significant at 1% in the *K*=18 model are also significant at 1% in the *K*=19 model. Finally, we observe that the coefficient of the 1997-year variable is much lower in both estimations than those of the other years.

The random-effects parameters are all significant in both estimations. In the hierarchical model, we assume that i) the fleet effect $\left(N_i\right)$ follows a negative binomial distribution with parameters $\left(\delta, \beta c\right)$; ii) the truck effect $\left(Z_{ij}\right)$, given N_i, is a negative binomial with parameters $\left(\delta + N_i, \frac{\beta c}{1+\beta c}\right)$; and iii) the time effect (θ_{ijt}), given Z_{ij}, follows a gamma distribution with parameters ($\delta^* + \beta^* Z_{ij}, c^*$). Because all these parameters are significant, the pricing formula of the hierarchical model will have to account for this additional information.

The estimation results of the Hausman model are presented in Table OA5.7 of Online appendix OA5. The Hausman model is suitable for estimating parameters with individual effects, but it cannot take into account common or fleet effects when individual observations belong to different firms with common characteristics that can affect accident distributions. Almost all coefficients of the Hausman model are also significant at 1%, including the two parameters, *a* and *b,* for the individual effects. We can use the BIC, the AIC, and the likelihood ratio test for comparison. We observe, in Table 2, that the hierarchical model, with *K*=19, performs better than the Hausman model for the different criteria presented in the table.[5]

We can compare the results of these two models with the gamma-Dirichlet model. The estimation results of the gamma-Dirichlet model are also presented in Table OA5.7, where we observe that the three random-effects parameters are significant. The significance of these

[5] This is also the case for *K*=18.

parameters means that the random effects associated with the fleets' unobservable risk (gamma with parameter κ), as well the random effects of trucks, including the drivers (Dirichlet with parameter ν), and the random time effects (Dirichlet with parameter ρ) significantly affect the distribution of truck accidents, even when we control for many observable characteristics.

We observe in Table 2 that the Hausman model performs better than the gamma-Dirichlet model, contrary to the results in Angers et al. (2018), where fleets with only one truck were not considered (see tables OA5.11 and OA5.12 for results with fleets of two trucks or more). The gamma-Dirichlet model seems to be penalized by the data containing many fleets that have only one truck, as in the current application (71% of fleet-years observations). We also obtain better estimation results with the hierarchical model than with the gamma-Dirichlet model for the different criteria presented in Table 2.

Table 1: Parameter estimation for the distribution of the number of annual truck accidents for the 1991–1998 period, for fleets with one truck or more, and trucks with two periods or more: hierarchical random-effects models with K=18 and K=19

	Hierarchical (K=19)		Hierarchical (K=18)	
Explanatory variable	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.1787*	0.0401	-3.1796*	0.0401
Number of years as a fleet	-0.0504*	0.0028	-0.0501*	0.0027
Sector of activity in 1998				
Other sectors	-0.1969	0.0967	-0.2004	0.0966
General public trucking	0.1022*	0.0273	0.0966*	0.0268
Bulk public trucking	Reference group			
Private trucking	0.0484	0.0203	0.0482	0.0204
Short-term rental firm	0.4795*	0.0497	0.4637*	0.0477
Size of fleet				
1	-0.0528*	0.0146	-0.0526*	0.0146
2	Reference group			
3	0.1205*	0.0156	0.1203*	0.0186
4 to 5	0.1766*	0.0185	0.1761*	0.0185
6 to 9	0.2381*	0.0195	0.2371*	0.0194
10 to 20	0.2824*	0.0204	0.2805*	0.0203
21 to 50	0.2447*	0.0243	0.2424*	0.0238

Explanatory variable	Hierarchical (*K*=19) Coefficient	Standard error	Hierarchical (*K*=18) Coefficient	Standard error
More than 50	0.2687*	0.0306	0.2735*	0.0249
Days in previous year	1.6494*	0.0245	1.6489*	0.0245
Violations				
Overload	0.0987*	0.0103	0.0990*	0.0103
Excess size	0.1588	0.0762	0.1587	0.0763
Poorly secured cargo	0.2435*	0.0331	0.2442*	0.0332
Not respecting service hours	0.1989*	0.0652	0.1999*	0.0653
No mechanical inspection	0.2171*	0.0259	0.2174*	0.0259
Other reasons	0.2175*	0.0695	0.2177*	0.0696
Type of vehicle use				
Commercial use	-0.1911*	0.0190	-0.1915*	0.0190
Other than bulk goods	-0.0807*	0.0230	-0.0799*	0.0229
Bulk goods	Reference group			
Type of fuel				
Diesel	Reference group			
Gas	-0.4317*	0.0120	-0.4325*	0.0120
Other	-0.3282*	0.0739	-0.3283*	0.0739
Number of cylinders				
1 to 5	0.2234*	0.0346	0.2268*	0.0346
6 to 7	0.3298*	0.0116	0.3299*	0.0116
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3451*	0.0184	-0.3447*	0.0184
2 axles (4,000 kg or more)	-0.3823*	0.0145	-0.3823*	0.0145
3 axles	-0.2995*	0.0142	-0.3014*	0.0141
4 axles	-0.2068*	0.0187	-0.2084*	0.0186
5 axles	-0.2518*	0.0163	-0.2524*	0.0162
6 axles or more	Reference group			
Number of violations				
Speeding	0.2198*	0.0092	0.2203*	0.0092
Suspended license	0.3958*	0.0331	0.3956*	0.0331
Running a red light	0.3828*	0.0207	0.3835*	0.0207
Ignoring a stop sign	0.4177*	0.0219	0.4178*	0.0219
Not wearing a seat belt	0.2139*	0.0247	0.2140*	0.0247
Observation period				
1991	-0.0015	0.0218	0.0003	0.0215
1992	-0.0318	0.0195	-0.0302	0.0193
1993	-0.0925*	0.0180	-0.0913*	0.0178
1994	-0.0230	0.0164	-0.0220	0.0163

Explanatory variable	Hierarchical (*K*=19) Coefficient	Standard error	Hierarchical (*K*=18) Coefficient	Standard error
1995	-0.0005	0.0151	0.0002	0.0150
1996	-0.0487*	0.0142	-0.0483*	0.0142
1997	-0.1568*	0.0140	-0.1566*	0.0140
1998	Reference group			
$\hat{\delta}$	0.7036*	0.0344	0.6725*	0.0314
$\widehat{\beta c}_0$	2.2025*	0.0616	2.1295*	0.0589
$\hat{\delta}^*$	2.6232*	0.2687	2.7240*	0.2728
$\hat{\beta}^*$	2.4959*	0.2687	2.6551*	0.2833
Number of observations	678,331		678,331	
Number of trucks	164,513		164,513	
Number of fleets	62,171		62,171	
Log likelihood	-265,353		-265,377	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.6877		0.6805	

* Significant at 1%.

Table 2: Fit statistics of the three models for fleets of one truck or more, and trucks with two periods or more

Statistics	Hausman model	Gamma-Dirichlet model	Hierarchical model *K*=19
Log likelihood	-269,077	-270,956	-265,353
BIC	538,758	542,530	531,337
AIC	538,244	542,004	530,800
Number of trucks	164,513	164,513	164,513
Number of observations	678,331	678,331	678,331
Number of firms	62,171	62,171	62,171
Number of parameters	45	46	47

The likelihood ratio test value of 7,448 is largely superior to the critical value of 9.21 at 1% when comparing the Hausman model to the hierarchical model with *K*=19. The likelihood ratio test value of 11,206 is largely superior to the critical value of 6.64 at 1% when comparing the gamma-Dirichlet model to the hierarchical model with *K*=19. The likelihood ratio test value of 3,758 is largely superior to the critical value of 6.64 at 1% when comparing the gamma-Dirichlet model to the Hausman 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 1, a premium must be a function of the observable characteristics of a vehicle and fleet, as well as a function of violations of the Highway Safety Code committed by drivers and fleet owners. These violations partially approximate the asymmetric information between the insurer and both the fleet owners and the drivers. As previously stated, this will not be enough, however, to obtain accurate pricing, because many unobservable actions of drivers and fleets owners may also affect the trucks' accident distribution. The premiums will have to be adjusted using the parameters of the random effects, to account for the impact of the unobservable characteristics of fleets and trucks as well as the unobservable behaviors of owners and drivers, and even for time not captured by the year variables. This form of rating makes it possible to visualize the impact (observable and unobservable) of the owners' and drivers' behaviors on the predicted accident rate, and consequently, on premiums for the next year. We now present the model's empirical pricing formula.

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 that 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:

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

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 contains the parameters estimated with hierarchical model (K = 19) for the 1991–1998 period, and the second set uses the parameters of hierarchical model (K = 18) estimated for the same period.

Figure 2 here

We can see from Figure 2 that the two distributions of the posterior expected number of accidents at period T+1 are very similar. The corresponding means and standard deviations (in parentheses) are, respectively, equal to 0.1368 (0.1048) for K = 18 and 0.1368 (0.1047) for K = 19. Both distributions differ significantly from the normal distribution.

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. Figure OA5.1 of Online appendix OA5 presents the distribution of the difference used for the paired t-test. The mean of the difference is not statistically different from zero at 1%. The t-test value is -0.66 and the p-value is equal to 0.51. From these results, we decided to proceed with K=19 for the out-of sample analysis. We should mention that the same analysis generated differences between means when comparing the results of the t-test with K=8 and K=10 and with K=13 and K=15. Results are presented in Online appendix OA5. See Table OA5.9 and Table OA5.10 and the corresponding figures.

6.2 Out-of-sample validation

To test the models' forecasting performance, our main estimation data are from the 1991–1996 observation period. The model is tested on the data from the validation year 1998. We also consider the 1991–1995 and 1991–1997 periods for estimation. The 1991–1995 period should be considered a robustness test period, and the results of the 1991–1997 period are presented in tables OA5.13 and OA5.14 for additional information, knowing that there is a mistake in the 1997 data. The database for the reoptimization of the hierarchical model in Table OA5.8 contains 491,792 observations for the 1991–1996 period, and 397,098 for the 1991–1995 period. The estimation results presented in Table OA5.8 are very stable between the two periods, so we should expect similar backtest results.

To compare the mean of the posterior expected number of accidents from estimations and the mean of the number of accidents in 1998, we run a *t*-test. The results are presented in Table 3 and Table 4. The mean of the posterior expected number of accidents is not statistically different from the mean of observed accidents in 1998, at a 1% level of significance, in Table 3 and Table 4, for all fleet sizes. Results in Table OA5.13 of the Online appendix OA5 perform less well for the estimation 1991–1997 period.

Table 3: *t*-test of the posterior expected number of accidents from the estimation of hierarchical model 91–96 and the observed numbers of accidents in 1998 for all fleets

	Hierarchical model 91–96			Data 1998			*t*-test	
	N trucks	Mean	Std	N trucks	Mean	Std	*t*-value	*p*-value
All fleets	132,868	0.1426	0.1129	77,651	0.1418	0.4109	0.58	0.5641
Size 1	33,965	0.1051	0.0796	25,230	0.1085	0.3581	-1.49	0.1375
Size 2	18,822	0.1164	0.3492	9,726	0.1092	0.3492	2.01	0.0448
Size 3	10,491	0.1273	0.0997	5,922	0.1348	0.3981	-1.42	0.1546
Sizes 4 to 5	12,669	0.1413	0.1089	7,321	0.1488	0.4521	-1.48	0.1399
Sizes 6 to 9	12,319	0.1605	0.1226	6,819	0.1684	0.4401	-1.45	0.1485
Sizes 10 to 20	13,785	0.1828	0.1384	7,754	0.1893	0.4751	-1.18	0.2375
Sizes 21 to 50	11,624	0.1828	0.1238	6,371	0.1874	0.4742	-0.76	0.4448
Sizes > 50	19,193	0.1794	0.1261	8,508	0.1776	0.4652	0.35	0.7249

Mean 91–96: Posterior expected number of accidents in year 1998 from estimations in Table OA5.8 (hierarchical model 91–96)
Data 1998: Observed mean of accidents by fleet size in 1998

Table 4: t-test of the posterior expected number of accidents from the estimation of hierarchical model 91–95 and the observed numbers of accidents in 1998 for all fleets

	Hierarchical model 91–95			Data 1998			t-test	
	N trucks	Mean	Std	N trucks	Mean	Std	t-value	p-value
All fleets	115,280	0.1420	0.1072	77,651	0.1418	0.4109	0.19	0.8524
Size 1	31,821	0.1043	0.0740	25,230	0.1085	0.3581	-1.83	0.0678
Size 2	16,086	0.1161	0.0821	9,726	0.1092	0.3492	1.92	0.0551
Size 3	9,105	0.1288	0.0956	5,922	0.1348	0.3981	-1.13	0.2566
Sizes 4 to 5	10,871	0.1422	0.0988	7,321	0.1488	0.4521	-1.30	0.1942
Sizes 6 to 9	10,278	0.1613	0.1176	6,819	0.1684	0.4401	-1.29	0.1983
Sizes 10 to 20	12,157	0.1841	0.1309	7,754	0.1893	0.4751	-1.95	0.3401
Sizes 21 to 50	9,605	0.1854	0.1163	6,371	0.1874	0.4742	-0.33	0.7412
Sizes > 50	15,357	0.1818	0.1270	8,508	0.1776	0.4652	0.82	0.4108

Mean 91–95: Posterior expected number of accidents in year 1998 from estimations in Table OA5.8 (hierarchical model 91–95)
Data 1998: Observed mean of accidents by fleet size in 1998

In the next section, we present premiums tables derived from the posterior expected number of accidents. To compute the premiums, we use the parameters of the two hierarchical models estimated with K= 19, as presented in Table OA5.8.

6.3 Application of the bonus-malus system

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

A premium for a truck at period T+1 can be affected by three possibilities, according to the theoretical model: 1) It has past experience and belongs to a fleet of trucks that has past experience; 2) It is a new truck, in a new fleet, meaning that there is no past experience for the truck or 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 and belonging to a fleet with past experience

Table 8 presents premiums for a truck, using all the information from the optimal estimations of the hierarchical model in Table OA5.8. From the hierarchical model and equation (16), the average estimated number of accidents is 0.1420 for the 91–95 estimation period and 0.1426 for the 91–96 estimation period. Both values are very close to the empirical mean of 0.1418, presented in Table OA5.4 for the year 1998. It decreases to less than 0.10 if the truck did not accumulate accidents in the past, but increases to more than 0.53 if it accumulated more than three accidents. The variations are similar for the two estimation periods.

Table 5: Premiums for a truck in 1998, as a function of the accumulated number of past accidents

Accumulated number of accidents over the period	Data	Hierarchical model 91–95			Data	Hierarchical model 91–96		
	1998	N trucks	$E[X_{ijT+1}\|X_{ikt}]$	Premium	1998	N trucks	$E[X_{ijT+1}\|X_{ikt}]$	Premium
0	0.0889	78,205	0.0993	$993	0.0918	88,828	0.0986	$986
1	0.1578	24,640	0.1791	$1,791	0.1584	28,602	0.1769	$1,769
2	0.2089	8,007	0.2769	$2,769	0.2100	9,657	0.2715	$2,715
3	0.2831	2,751	0.3828	$3,828	0.2633	3,471	0.3708	$3,708
More than 3	0.4665	1,677	0.5518	$5,518	0.4597	2,310	0.5397	$5,397
Total	0.1237	115,280	0.1420	$1,420	0.1300	132,868	0.1426	$1,426

Premium for the posterior expected number of accidents in 1998, from estimations in Table OA5.8, equation (16), and the conditional average cost of claims of $10,000. Of the 115,280 trucks in the 91–95 period, 39,289 are still present in 1998. The 1998 data are their average number of accidents in 1998, given their past experience in 91–95. The respective numbers for the 91–96 period are 50,002 trucks in 1998, from the 132,868 trucks in 91–96.

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 $\mathrm{E}[\theta_{i,j,T+1}|x_{i,k,t}] = \mathrm{E}[\theta_{i,j,t}] = 1$.

The posterior expected number of accidents for a new truck in a new fleet will be

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

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 the new fleet, we represent the truck with a vehicle that is representative of the population, i.e., λ. However, even if this is a new truck, we know its observable characteristics, such as the type of gasoline it uses, etc. We can thus obtain various scenarios for λ, according to the characteristics of the truck and the fleet. Table 6 presents the premiums for a new truck from a new fleet. The average estimated numbers of accidents are slighly different from the values obtained in Table OA5.4. Data by fleet size are also documented. The premiums fall if the truck belongs to a fleet of a lower size, as observed with the estimated parameters in Table OA5.8.

Table 6: Premiums for a new truck in a new fleet in 1998

	Data 1998	Hierarchical model 91–95			Hierarchical model 91–96		
	Mean	N trucks	λ	Premium	N trucks	λ	Premium
Size of fleet							
1	0.1085	31,821	0.1041	$1,041	33,965	0.1067	$1,067
2	0.1092	16,086	0.1146	$1,146	18,822	0.1174	$1,174
3	0.1348	9,105	0.1263	$1,263	10,491	0.1276	$1,276
4 to 5	0.1488	10,871	0.1422	$1,422	12,669	0.1430	$1,430
6 to 9	0.1684	10,278	0.1599	$1,599	12,319	0.1630	$1,630
10 to 20	0.1893	12,157	0.1793	$1,793	13,785	0.1833	$1,833
21 to 50	0.1874	9,605	0.1789	$1,789	11,624	0.1849	$1,849
More than 50	0.1776	15,357	0.1714	$1,714	19,193	0.1775	$1,775
Total	0.1418	115,280	0.1390	$1,390	132,868	0.1436	$1,436

Premium for the posterior expected number of accidents, from estimations in Table OA5.8, equation (17), and the conditional average cost of claims of $10,000.

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

If the fleet existed previously, there may not be any past accidents recorded for the new truck, but there are past accidents for the 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

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

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 7 presents the premiums for a new truck in an existing fleet. The average estimated numbers of accidents are very similar to those obtained in Table 5.

Table 7: Premiums for a new truck in 1998 in an existing fleet

	Data 1998	Hierarchical model 91–95			Hierarchical model 91–96		
	Mean	N trucks	$E[X_{ijT+1}\|X_{ikt}]$	Premium	N trucks	$E[X_{ijT+1}\|X_{ikt}]$	Premium
Size of fleet							
1	0.1085	31,821	0.1043	$1,043	33,965	0.1051	$1,051
2	0.1092	16,086	0.1161	$1,161	18,822	0.1164	$1,164
3	0.1348	9,105	0.1288	$1,288	10,491	0.1273	$1,273
4 to 5	0.1488	10,871	0.1422	$1,422	12,669	0.1413	$1,413
6 to 9	0.1684	10,278	0.1613	$1,613	12,319	0.1605	$1,605
10 to 20	0.1893	12,157	0.1841	$1,841	13,785	0.1828	$1,828
21 to 50	0.1874	9,605	0.1854	$1,854	11,624	0.1828	$1,828
More than 50	0.1776	15,357	0.1818	$1,818	19,193	0.1794	$1,794
Total	0.1418	115,280	0.1420	$1,420	132,868	0.1426	$1,426

Posterior expected number of accidents from the estimations in Table OA5.8, equation (18), and the conditional average cost of claims of $10,000.

Conclusion

In this paper, we derive a new hierarchical count-data model with random effects to estimate and forecast truck accidents within a fleet. We develop a closed-form formula for the ratemaking of a bonus-malus scheme that considers past accidents, past road safety offences of both the drivers and the fleet owners, and the characteristics of the trucks and fleets.

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. We should keep in mind that, in many applications, such as insurance, banking, education, and marketing, the aim of the econometric model is not only to explain the data and identify the most important explanatory variables, but also to forecast the response variables. Our choice of random effects is related to the fact that most fleets in our data are of moderate size and were observed over a small number of years. In this case, a fixed (fleet) effect model could be generally inconsistent, due to the incidental parameter problem. On the other hand, a potential downside of the random-effects model is that it might also suffer from endogeneity bias, if the independence between the random effects $\theta_{i,j,t}$ and the covariate vector $Y_{i,j,t}$ is not satisfied. In our nonlinear panel-data model, however, this issue can be partially mitigated by considering more flexible, possibly nonparametric, models for the marginal conditional expectation $\lambda_{i,j,t} = E\left[X_{i,j,t} \middle| Y_{i,j,t}\right]$, beyond the standard exponential linear specification.[6] Consequently, we believe that the random-effects model is more appropriate for our application, especially given its tractable posterior predictive formula, which constitutes the very foundation of actuarial bonus-malus pricing.

Moreover, in our specification, the effect of time is fleet- and truck-specific, since it only appears at the third level. This structure, which implicitly puts less emphasis on intertemporal variation than on intertruck and interfleet variation, is motivated by the principal's (insurer, regulator) desire to incentivize fleet managers and truck drivers to comply with road safety in future years. It might also be (at least econometrically) interesting to consider other hierarchical structures. For instance, one can continue to have fleet effects at the first level, but fleet/time effects

[6] Alternatively, Hausman et al. (1984) also propose a conditional maximum likelihood estimation approach that is robust to the potential existence of a correlation between the random effects and the covariates. Nevertheless, this method does not address the endogeneity issue when it comes to forecasting. See Chamberlain (1980) on endogeneity in a general context, not specific to count models.

at the second level, and fleet/time/truck effects at the third level. This latter specification may be more suitable if, empirically, there is significant temporal fluctuation in the fleets' average risk.

The insurance policy considered in the paper is renewed annually; hence, all the data, including the claims counts, are only observed at a quite low (annual) frequency. Recently, telematic insurance has become increasingly popular for individual vehicles. The insurer can now collect high-frequency count data for vehicles. To our knowledge, telematic data has not yet been introduced for fleet insurance, but because of its tractability, the hierarchical count-data model we propose here would be a serious contender once such data become available. We leave these issues for future research.

References

Abowd, J. & Kramarz F. (1999). Econometric analyses of linked employer-employee data. *Labour Economics, 6*, 53-74.

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

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

Böckenholt, U. (1998). Mixed INAR (1) Poisson regression models: Analyzing heterogeneity and serial dependencies in longitudinal count data. *Journal of Econometrics, 89*, 317-338.

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.

Chamberlain, G.E. (2013). Predictive effects of teachers and schools on test scores, college attendance, and earnings. *Proceedings of the National Academy of Sciences, 110*, 17176-17182.

Chamberlain, G. (1980). Analysis of covariance with qualitative data. *Review of Economic Studies*, *47*, 225-238.

Chib, S. & Winkelmann, R. (2001). Markov chain Monte Carlo analysis of correlated count data. *Journal of Business & Economic Statistics, 19*, 428-435.

Connolly, J.M. (2016). The impact of local politics on the principal-agent relationship between council and managers in municipal government. *Journal of Public Administration Research and Theory*, *27*, 253-268.

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

Dionne, G. & Dostie, B. (2007) &. New evidence on the determinants of absenteeism using linked employer-employee data. *ILR Review, 61*, 108-120.

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: Dionne, G. and Laberge-Nadeau, C. (Eds.), *Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation*, 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.

Englin, J. & Shonkwiler, J.S. (1995). Estimating social welfare using count data models: an application to long-run recreation demand under conditions of endogenous stratification and truncation. *Review of Economics and Statistics, 77,* 104-112.

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.

Freeman, R.B. & Viarengo, M. (2014). School and family effects on educational outcomes across countries. *Economic Policy*, *29*, 395-446.

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

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

Heston, S.L. & Rossi, A.G. (2017). A spanning series approach to options. *Review of Asset Pricing Studies*, *7*, 2-42.

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

Holmstrom, B. & Milgrom, P. (1987). Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, *55*, 303-328.

Hubert, E. (2020). Continuous-time incentives in hierarchies. Working paper, LAMA, Université Gustave Eiffel.

Kim, J. & Marschke, G. (2005). Labor mobility of scientists, technological diffusion, and the firm's patenting decision. *RAND Journal of Economics, 36*, 298-317.

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

Lee, W., Kim, J., & Ahn, J.Y. (2020). The Poisson random effect model for experience ratemaking: Limitations and alternative solutions. *Insurance: Mathematics and Economics*, *91*, 26-36.

Ludwig, M., Merode, F.V., & Groot, V. (2010). Principal agent relationships and the efficiency of hospitals. *The European Journal of Health Economics*, *11*, 291-304.

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* (2nd ed.). Springer, New York, 471-586.

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

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

Sung, J. (2015). Pay for performance under hierarchical contracting. *Math Finan Econ* 9, 195-213.

Winkelmann, R. (2004). Health care reform and the number of doctor visits—an econometric analysis. *Journal of Applied Econometrics*, *19*, 455-472.

Zeger, S.L. & Karim, M.R. (1991). Generalized linear models with random effects: A Gibbs sampling approach. *Journal of the American Statistical Association*, *86*, 79-86.

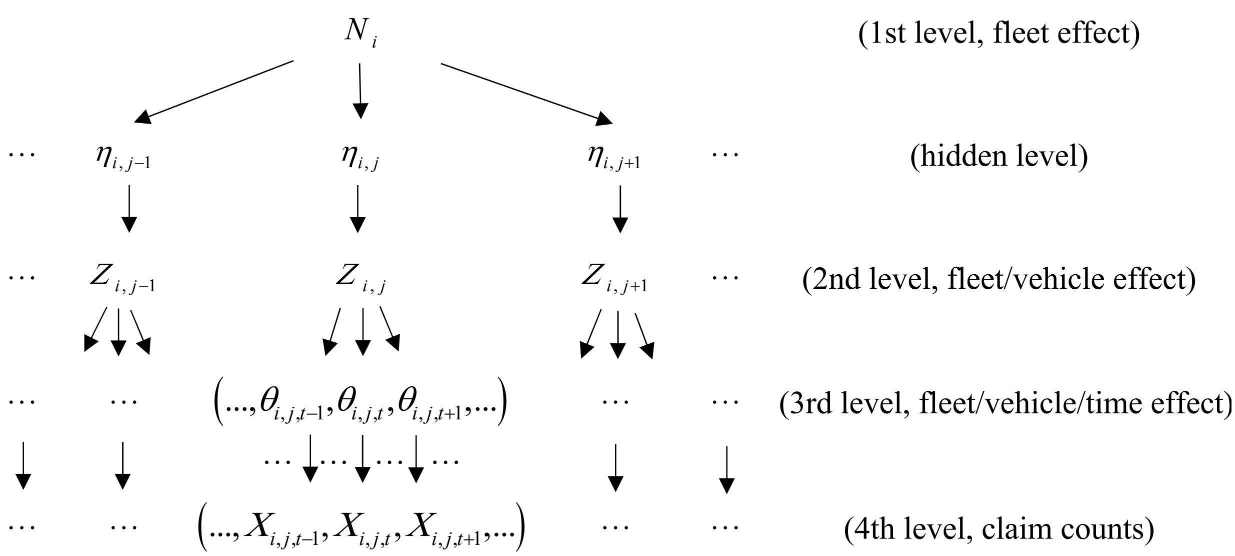

Figure 1: Chain rule of the hierarchical model

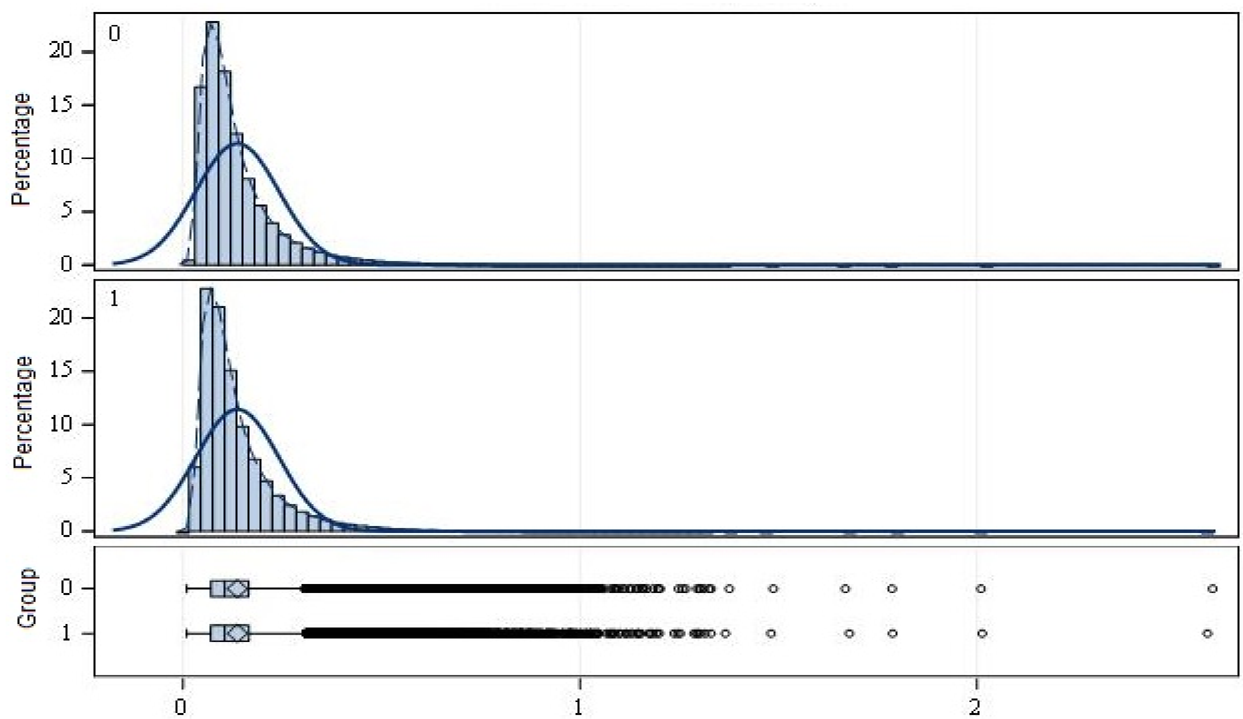

Figure 2: Posterior expected number of accidents at period T+1, obtained from the estimation of hierarchical model K=18 (at top) and hierarchical model K=19 (at bottom). The continuous density function is from the normal distribution.

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

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

17 October 2022

Online appendices

OA1 Hierarchical principal-agent model: principal-agent relationship in a teams model with an intermediary

Fleet model

Our contribution is related to the problem of moral hazard in teams (Holmstrom, 1982), although our environment is not exactly similar since we add an intermediary between the principal and the agents. Moral hazard is an information problem of inducing agents to produce optimal effort when their action cannot be observed and contracted directly (Laffont and Martimort, 2001). Moral hazard in teams applies to several agents, and hierarchical moral hazard is when an intermediary between the principal and the agents is added (Sung, 2015; Hubert, 2020).

This model is an extension of the Holmstrom and Milgrom (1987) model of interaction between economic agents with asymmetric information in a hierarchical framework. This extension of the multi-agent model was first proposed by Sung (2015) in a one-period principal-agent framework with moral hazard. In our application, the agents are the truck drivers. The principal is the insurer, and the fleet owner is the intermediary who can influence the agents' safety output through monitoring and own actions. The actions of both the agent and the intermediary to improve road safety are private information, meaning that they cannot be observed directly by the insurer. The intermediary also does not observe the agent's actions.

Moral hazard often occurs when there is uncertainty in output realization, as in road accidents. The actions of agents or intermediaries who cheat by committing road safety offences cannot be identified when observing accidents, which is the only, imperfect, observable indicator of road safety output in this type of model. An accident can occur because of bad driving or bad luck, and observing only accidents does not allow for conclusions about the cause of the accidents, especially under no-fault insurance, without a full accident investigation, as in our environment.

For simplicity, the model is limited to the relationships between the parties with respect to the contracting of insurance for a fleet of vehicles and its links to road safety. The other responsibilities of the fleet owner, such as hiring drivers, improving fleet productivity, or being competitive in the transportation of goods, are considered independent.

The insurer writes an insurance contract for the trucks of the fleet and sells it to the fleet owner. Accident distributions can be affected independently by the agents and the intermediary. In our application, unobservable actions are road safety activities that can affect the mean of the accident distributions, in a first-order manner. It is less costly for insurers to let the fleet owners monitor the drivers.

In our application, the principal is risk-neutral, the intermediary is risk-averse, and the agents are risk-averse. This explains insurance demand. The agents can generate accidents, which are the agent's outputs that are observable by the intermediary. Accidents are costly to the intermediary, so they buy insurance coverage by paying premiums to the insurer. The principal observes only accidents. The insurer sets the aggregate premium and reimburses the aggregate accident costs to the fleet owner, as per the insurance contract. The intermediary observes the truck accidents and pays a salary to the truck drivers that is a decreasing function of accidents. The intermediary reports all accident costs to the insurer and pays the premiums, which are increasing functions of accidents.

In a similar framework applied to finance, Sung (2015) shows that linear contracts are optimal in a one-period discrete-time model where the intermediary cannot affect the volatility of the returns distribution. This is a kind of coinsurance contract in an insurance application.

In the insurance literature, multiperiod contracts have been implemented to observe proxies over time, for past effort levels, such as past accidents and past offences, in order to improve resource allocation under moral hazard. We assume that the intermediary uses such information to compute the individual premiums of the trucks and to set the corresponding salaries of the truck drivers.

We do not pretend that this form of contracting is the optimal one. More theoretical research would be necessary to obtain the optimal form of both the insurance and the labor contracts, which falls beyond the purview of this article. We add the safety efforts of both the intermediaries and the drivers in the regression components as control variables, but they are not contract decision variables in this model. Only accidents are contract decision variables for setting the insurance premiums.

Other applications

Children's academic success: The principal is the school director, the agents are the students, and the intermediaries are the parents. The students' effort, which affects their success rate (number of As per period) is not perfectly observed by the parents or the school director. The effort of the parents is not perfectly observed by the school director, and their effort can affect the students' success rate.

Absence from work: The principal is the owner of the company, the agent is the worker, and the intermediary is the supervisor.

Surgical malpractice: The principal is the insurer, the agent is the surgeon, and the intermediary is the owner of the clinic or hospital.

Finance: The principal is the investor, the agent is the trader, and the intermediary is the broker.

Business firm: The principal is the stock owner, the agents are the workers, and the intermediary is the CEO.

OA2 The Poisson-gamma conjugacy

The following proposition reviews the main results of the Poisson-gamma conjugacy.

Proposition 2: Poisson-gamma conjugacy

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 ν on N and the Lebesgue measure* λ^+ *on* R^+*):*

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

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:* $\gamma\left(\delta,\frac{c}{1-\beta c}\right)$.

□

OA3 Gamma-Dirichlet model

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

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

To simultaneously take into account of both the firm effect and the time effect, suppose that $\lambda_{ijt} = \gamma_{ijt}\left(\alpha_i \theta_{(i)j} \eta_{(ij)t}\right)$ with $\gamma_{ijt} = e^{X_{ijt}\beta}$. The vector $X_{ijt} = \left(x_{ijt1}, \cdots, x_{ijtp}\right)$ represents the *p* characteristics of truck *j* of fleet *i* 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_i be the random effects associated with fleet *i* (i.e. the risk or non-observable characteristics attributable to the fleet), whereas $\theta_{(i)j}$ is the random effects of truck *j* of fleet *i* where $\sum_{j=1}^{s_i} \theta_{(i)j} = 1$ where s_i is the number of vehicles in fleet *i*. Finally, $\eta_{(ij)t}$ is the random effects of period *t* of truck *j* of fleet *i* such as $\sum_{t=1}^{T_j} \eta_{(ij)t} = 1$ where T_j is the number of periods for truck *j*.

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

$$P\left(Y_{i11},\cdots,Y_{is_iT_{s_i}} \mid \underset{\sim}{\eta}_{(ij)}\right)$$

$$=\frac{\left[\Gamma\left(S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1}\right)\right]\left[\left(\kappa^{-1}\right)^{\sum_{j=1}^{s_i}T_j\kappa^{-1}}\right]\left[\Gamma\left(\sum_{j=1}^{s_i}\nu_{(i)j}\right)\right]\left[\prod_{j=1}^{s_i}\prod_{t=1}^{T_j}\left(\gamma_{ijt}\eta_{(ij)t}\right)^{y_{ijt}}\right]}{\left[\prod_{j=1}^{s_i}\prod_{t=1}^{T_j}\Gamma\left(y_{ijt}+1\right)\right]\left[\Gamma\left(\sum_{j=1}^{s_i}T_j\kappa^{-1}\right)\right]\left[\prod_{j=1}^{s_i}\Gamma\left(\nu_{(i)j}\right)\right]}\int\cdots\int\frac{\prod_{j=1}^{s_i}\left(\theta_{(i)j}\right)^{S_j+\nu_{(i)j}-1}}{\left(\kappa^{-1}+\sum_{j=1}^{s_i}\theta_{(i)j}\sum_{t=1}^{T_j}\gamma_{ijt}\eta_{(ij)t}\right)^{S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1}}}\,d\underset{\sim}{\theta}_{(i)}.$$

(A2)

By supposing that $\gamma_{ijt}=\overline{\gamma}_{ij}\forall t=1,\cdots,T_j$ where $\overline{\gamma}_{ij}=\frac{1}{T_j}\sum_{j=1}^{T_j}\gamma_{ijt}$, 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 $\overline{\gamma}_{g_1}=\frac{\sum_{j=1}^{g_1}\overline{\gamma}_{ij}}{g_1}$, and $G_2=g_1+1,\cdots,s_i$, as the set of all vehicles of the second group with $\overline{\gamma}_{g_2}=\frac{\sum_{j=g_1+1}^{s_i}\overline{\gamma}_{ij}}{g_2}$, the integral of equation (A2) thus becomes:

$$\int\cdots\int\frac{\left[\prod_{j=1}^{g_1}\left(\theta_{(i)j}\right)^{S_j+\nu_{(i)j}-1}\right]\left[\prod_{j=g_1+1}^{s_i}\left(\theta_{(i)j}\right)^{S_j+\nu_{(i)j}-1}\right]}{\left(\kappa^{-1}+\overline{\gamma}_{g_1}\sum_{j=1}^{g_1}\theta_{(i)j}+\overline{\gamma}_{g_2}\sum_{j=g_1+1}^{s_i}\theta_{(i)j}\right)^{S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1}}}\,d\underset{\sim}{\theta}_{(i)}. \quad \text{(A3)}$$

By integration, one obtains:

$$
= \frac{\prod_{j=1}^{s_i} \Gamma\left(S_j + \nu_{(i)j}\right)}{\left[\Gamma\left(\sum_{j=1}^{s_i} S_j + \nu_{(i)j}\right)\right]\left[\left(\kappa^{-1} + \overline{\gamma}_{g_2}\right)^{S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}}\right]} {}_2F_1\left(\sum_{j=1}^{g_1} S_j + \nu_{(i)j}, S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}, \sum_{j=1}^{s_i} S_j + \nu_{(i)j}, \left(\frac{\overline{\gamma}_{g_2} - \overline{\gamma}_{g_1}}{\kappa^{-1} + \overline{\gamma}_{g_2}}\right)\right).
$$

(A4)

Thus, by replacing the integral in equation (A3) with its value given in (A4) the following approximation for $P\left(Y_{i11}, \cdots, Y_{is_iTs_i} \mid \underset{\sim}{\eta}_{(ij)}\right)$ in (A2) is equal to:

$$
\frac{\Gamma\left(S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}\right)\left(\kappa^{-1}\right)^{\sum_{j=1}^{s_i} T_j \kappa^{-1}} \Gamma\left(\sum_{j=1}^{s_i} \nu_{(i)j}\right) \prod_{j=1}^{s_i} \prod_{t=1}^{T_j} \left(\gamma_{ijt} \eta_{(ij)t}\right)^{y_{ijt}}}{\prod_{j=1}^{s_i} \prod_{t=1}^{T_j} \Gamma\left(y_{ijt} + 1\right) \Gamma\left(\sum_{j=1}^{s_i} T_j \kappa^{-1}\right) \prod_{j=1}^{s_i} \Gamma\left(\nu_{(i)j}\right)} \left(\frac{\prod_{j=1}^{s_i} \Gamma\left(S_j + \nu_{(i)j}\right)}{\Gamma\left(\sum_{j=1}^{s_i} S_j + \nu_{(i)j}\right)}\right) \frac{1}{\left(\kappa^{-1} + \overline{\gamma}_{g_2}\right)^{S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}}}
$$

$$
\times {}_2F_1\left(\sum_{j=1}^{g_1} S_j + \nu_{(i)j}, S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}, \sum_{j=1}^{s_i} S_j + \nu_{(i)j}, \left(\frac{\overline{\gamma}_{g_2} - \overline{\gamma}_{g_1}}{\kappa^{-1} + \overline{\gamma}_{g_2}}\right)\right) \qquad \text{(A5)}
$$

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

$$
1 + \sum_{\ell=1}^{\infty} \left[\frac{\left(\sum_{j=1}^{g_1} S_j + \nu_{(i)j}\right)^{[\ell]} \left(S_0 + \sum_{j=1}^{s_i} T_j \kappa^{-1}\right)^{[\ell]}}{\left(\sum_{j=1}^{s_i} S_j + \nu_{(i)j}\right)^{[\ell]}} \frac{\left(\frac{\overline{\gamma}_{g_2} - \overline{\gamma}_{g_1}}{\kappa^{-1} + \overline{\gamma}_{g_2}}\right)^{\ell}}{\ell!} \right],
$$

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

$$\left[\prod_{j=1}^{s_i}\prod_{t=1}^{T_j}\frac{\left(\gamma_{ijt}\right)^{y_{ijt}}}{\Gamma\left(y_{ijt}+1\right)}\right]\left[\frac{\Gamma\left(S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1}\right)}{\Gamma\left(\sum_{j=1}^{s_i}T_j\kappa^{-1}\right)}\right]\left[\frac{\left(\kappa^{-1}\right)^{\sum_{j=1}^{s_i}T_j\kappa^{-1}}}{\left(\kappa^{-1}+\overline{\gamma}_{g_2}\right)^{S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1}}}\right]\left[\frac{\Gamma\left(\sum_{j=1}^{s_i}\nu_{(i)j}\right)}{\Gamma\left(\sum_{j=1}^{s_i}S_j+\nu_{(i)j}\right)}\right]\left[\frac{\prod_{j=1}^{s_i}\Gamma\left(S_j+\nu_{(i)j}\right)}{\prod_{j=1}^{s_i}\Gamma\left(\nu_{(i)j}\right)}\right]$$

$$\times\left[\frac{\prod_{j=1}^{s_i}\prod_{t=1}^{T_j}\Gamma\left(y_{ijt}+\rho_{(ij)t}\right)}{\prod_{j=1}^{s_i}\prod_{t=1}^{T_j}\Gamma\left(\rho_{(ij)t}\right)}\right]\left[\frac{\prod_{j=1}^{s_i}\Gamma\left(\sum_{t=1}^{T_j}\rho_{(ij)t}\right)}{\prod_{j=1}^{s_i}\Gamma\left(S_j+\sum_{t=1}^{T_j}\rho_{(ij)t}\right)}\right]\times{}_2F_1\left(\sum_{j=1}^{g_1}\left(S_j+\nu_{(i)j}\right),S_0+\sum_{j=1}^{s_i}T_j\kappa^{-1},\sum_{j=1}^{s_i}\left(S_j+\nu_{(i)j}\right),\left(\frac{\overline{\gamma}_{g_2}-\overline{\gamma}_{g_1}}{\kappa^{-1}+\overline{\gamma}_{g_2}}\right)\right),$$

(A6)

where:

s_i is the number of vehicles in fleet *i*.

T_j is the number of periods for truck *j*.

$S_0=\sum_{j=1}^{s_i}\sum_{t=1}^{T_j}y_{ijt}=\sum_{j=1}^{s_i}S_j$ where $S_j=\sum_{t=1}^{T_j}y_{ijt}$.

Letting $\nu_{(i)j}=\nu,\forall j$ and $\rho_{(ij)t}=\rho,\forall t$, we can use the maximum likelihood method to estimate the unknown parameters, ν,κ,ρ and β of the log likelihood function of the model.

Weakenesses of the gamma-Dirichlet model. The gamma-Dirichlet model has two weakenesses. First, the above formulas involve approximations. Second and more importantly, the model suffers from a self-consistency issue. More precisely, the identification condition $\sum_{j=1}^{s_i}\theta_{(i)j}=1$ depends on the number of vehicles s_i. Thus when the size of a fleet changes, this condition is no longer satisfied. In other words, for forecasting purpose, the new model is not consistent with the initial one. Similarly, condition $\sum_{t=1}^{T_j}\eta_{(ij)t}=1$ depends on the number of years T_j . Thus, the model used in the next year is incompatible with the model currently in place.

OA4 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 explanatory 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:* 8 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 activity*: 5 dichotomous variables have been created for vehicles transporting goods:

  sect_14 = 1 if the main sector of activity is transporting passengers; **(Other sectors)**

  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.

- 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. **(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. **(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. **(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 respecting 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. (**Other reasons)**

Variables concerning the vehicle

- *Vehicle's number of cylinders:* 3 dichotomous variables have been created:

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.

- *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.

- *Maximum number of axles:* 6 dichotomous variables have been created:

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.

- *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.

- Five variables have been created to measure the number of convictions per vehicle cumulated 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. **(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. **(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. **(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)**

OA5 Other statistics and results

Other statistics

Table OA5.1: Fleet size, number of years of follow-up by fleet size, and number of years of follow-up by truck in a given fleet size

	Number of years of follow-up by fleet size			Number of years of follow-up by truck in a given fleet size		
Fleet size	Number of fleets	Mean	Median	Number of trucks	Mean	Median
1	38,272	4.06	3	38,272	4.06	3
2	11,628	5.96	6	23,256	3.88	3
3	4,396	6.47	7	13,188	4.04	3
4 to 5	3,620	6.80	8	15,810	4.20	4
6 to 9	2,256	7.00	8	15,948	4.28	4
10 to 20	1,302	7.06	8	17,574	4.30	4
21 to 50	496	7.20	8	15,049	4.25	4
More than 50	201	7.04	8	25,416	3.92	3
Total	62,171	4.95	5	164,513	4.13	4

Table OA5.2: Number of years of follow-up of the fleet and of the truck

Number of years of follow-up	Fleet		Truck	
	N	%	N	%
2	13,059	21.00	45,308	27.54
3	9,349	15.04	34,495	20.97
4	7,533	12.12	26,112	15.87
5	6,041	9.72	17,740	10.78
6	5,340	8.59	13,356	8.12
7	5,542	8.91	9,070	5.51
8	15,307	24.62	18,432	11.20
Total	62,171	100.00	164,513	100.00

Table OA5.3: Size of fleet distribution (in %) by year

Size of fleet	% by year								% total of fleet-years
	1991	1992	1993	1994	1995	1996	1997	1998	
1	72.76	71.38	70.90	70.53	70.37	70.25	70.56	70.95	70.91
2	13.05	13.53	13.86	14.10	13.95	13.99	13.79	13.67	13.76
3	5.35	5.59	5.70	5.75	5.69	5.70	5.64	5.55	5.63
4 to 5	4.15	4.47	4.50	4.47	4.71	4.66	4.72	4.72	4.56
6 to 9	2.49	2.64	2.62	2.72	2.78	2.87	2.76	2.69	2.70
10 to 20	1.48	1.60	1.64	1.65	1.66	1.65	1.71	1.64	1.63
21 to 50	0.53	0.58	0.57	0.54	0.59	0.62	0.60	0.58	0.58
More than 50	0.18	0.21	0.22	0.24	0.25	0.26	0.24	0.20	0.23
Number of fleets	31,793	38,236	39,128	39,882	40,688	41,214	41,289	35,562	307,792

Table OA5.4: Truck accident distribution by year of observation

Annual truck accidents	Year of observation								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
0	87.23	87.61	88.50	88.03	87.82	88.40	89.60	87.81	88.17
1	10.95	10.67	10.02	10.34	10.47	10.06	9.18	10.53	10.24
2	1.53	1.45	1.26	1.38	1.42	1.28	1.05	1.38	1.34
3	0.23	0.22	0.19	0.20	0.24	0.21	0.15	0.23	0.21
4 and more	0.05	0.05	0.04	0.05	0.05	0.05	0.03	0.05	0.05
Number of trucks	66,193	83,230	86,090	88,152	91,260	93,357	92,398	77,651	678,331
Mean accident rate	0.1495	0.1443	0.1325	0.1390	0.1423	0.1345	0.1185	0.1418	0.1372

Table OA5.5: Average truck accidents per fleet size by year

Size of fleet	% by year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
1	0.1115	0.1065	0.0991	0.1046	0.1055	0.0976	0.0884	0.1085	0.1024
2	0.2628	0.2453	0.2204	0.2207	0.2224	0.2156	0.1805	0.2184	0.2216
3	0.4365	0.4186	0.3807	0.4002	0.4135	0.3722	0.3149	0.4043	0.3905
4 to 5	0.6689	0.6865	0.6016	0.6272	0.6477	0.5856	0.5506	0.6490	0.6242
6 to 9	1.3914	1.2232	1.0898	1.1335	1.1830	1.0990	1.0026	1.1996	1.1551
10 to 20	2.6730	2.6144	2.3744	2.5099	2.4541	2.4824	2.0767	2.5223	2.4501
21 to 50	5.9176	5.3818	5.0448	5.3565	5.8875	5.2461	4.8618	5.7681	5.4094

Size of fleet	% by year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
More than 50	21.1034	21.3293	20.8023	21.2021	20.9700	20.3962	17.5859	21.2817	20.4971
Average truck accidents per fleet	0.3109	0.3141	0.2915	0.3071	0.3192	0.3048	0.2652	0.3095	0.3023

Table OA5.6: Average truck accidents rates according to the driver's and fleet owners' violations committed the previous year

Violations committed by the driver the previous year	Year								Total
	1991	1992	1993	1994	1995	1996	1997	1998	
Speeding									
0	0.1438	0.1399	0.1268	0.1325	0.1348	0.1266	0.1104	0.1335	0.1305
1	0.2850	0.2402	0.2480	0.2524	0.2438	0.2406	0.2010	0.2504	0.2405
2	0.2512	0.3067	0.3344	0.3691	0.3752	0.3041	0.2950	0.3095	0.3200
3 and more	0.4750	0.5000	0.2653	0.4154	0.4766	0.4452	0.3707	0.4650	0.4269
Suspended license									
0	0.1493	0.1441	0.1320	0.1381	0.1411	0.1337	0.1179	0.1412	0.1366
1 and more	0.6667	0.3750	0.3672	0.3680	0.3234	0.3056	0.2622	0.3106	0.3207
Running a red light									
0	0.1473	0.1426	0.1310	0.1369	0.1398	0.1322	0.1167	0.1391	0.1351
1	0.2737	0.2771	0.2536	0.2985	0.3244	0.2993	0.2771	0.3555	0.2954
2 and more	0.4828	0.5417	0.5385	0.4762	0.6429	0.5625	0.2778	0.5625	0.5206
Ignoring a stop sign or police signal									
0	0.1472	0.1426	0.1310	0.1373	0.1399	0.1326	0.1172	0.1401	0.1354
1	0.3063	0.2966	0.2553	0.2782	0.3342	0.2792	0.2299	0.2748	0.2815
2 and more	0.4074	0.5000	0.2609	0.5625	0.4167	0.6286	0.2083	0.7200	0.4694
Not wearing a seat belt									
0	0.1487	0.1438	0.1317	0.1385	0.1415	0.1337	0.1174	0.1415	0.1365
1	0.1991	0.1933	0.2095	0.1758	0.2037	0.1941	0.1888	0.1855	0.1943
2 and more	0.3061	0.3404	0.2459	0.3553	0.2188	0.2632	0.2889	0.2000	0.2762
	Year								Total

Violations committed by the fleet owner the previous year	1991	1992	1993	1994	1995	1996	1997	1998	
Overload									
0	0.1443	0.1395	0.1282	0.1352	0.1373	0.1288	0.1147	0.1381	0.1327
1	0.2413	0.2698	0.2329	0.2365	0.2365	0.2353	0.1876	0.2584	0.2339
2 and more	0.3333	0.3067	0.3498	0.2838	0.3172	0.3051	0.2121	0.3878	0.3008
Excessive size									
0	0.1492	0.1443	0.1324	0.1389	0.1423	0.1344	0.1184	0.1417	0.1371
1 and more	0.2921	0.1346	0.2368	0.2174	0.1677	0.1962	0.1902	0.2308	0.2044
Poorly secured cargo									
0	0.1485	0.1436	0.1436	0.1319	0.1383	0.1340	0.1180	0.1412	0.1365
1 and more	0.3167	0.3156	0.3156	0.2529	0.2651	0.2599	0.2564	0.3667	0.2777
Not respecting service hours									
0	0.1492	0.1443	0.1323	0.1389	0.1420	0.1344	0.1183	0.1414	0.1370
1 and more	0.6000	0.2727	0.5854	0.1837	0.3680	0.2623	0.3506	0.3761	0.3502
No mechanical inspection									
0	0.1488	0.1435	0.1309	0.1379	0.1407	0.1340	0.1178	0.1412	0.1363
1 and more	0.2660	0.3280	0.2050	0.2233	0.2779	0.2059	0.2098	0.2588	0.2344

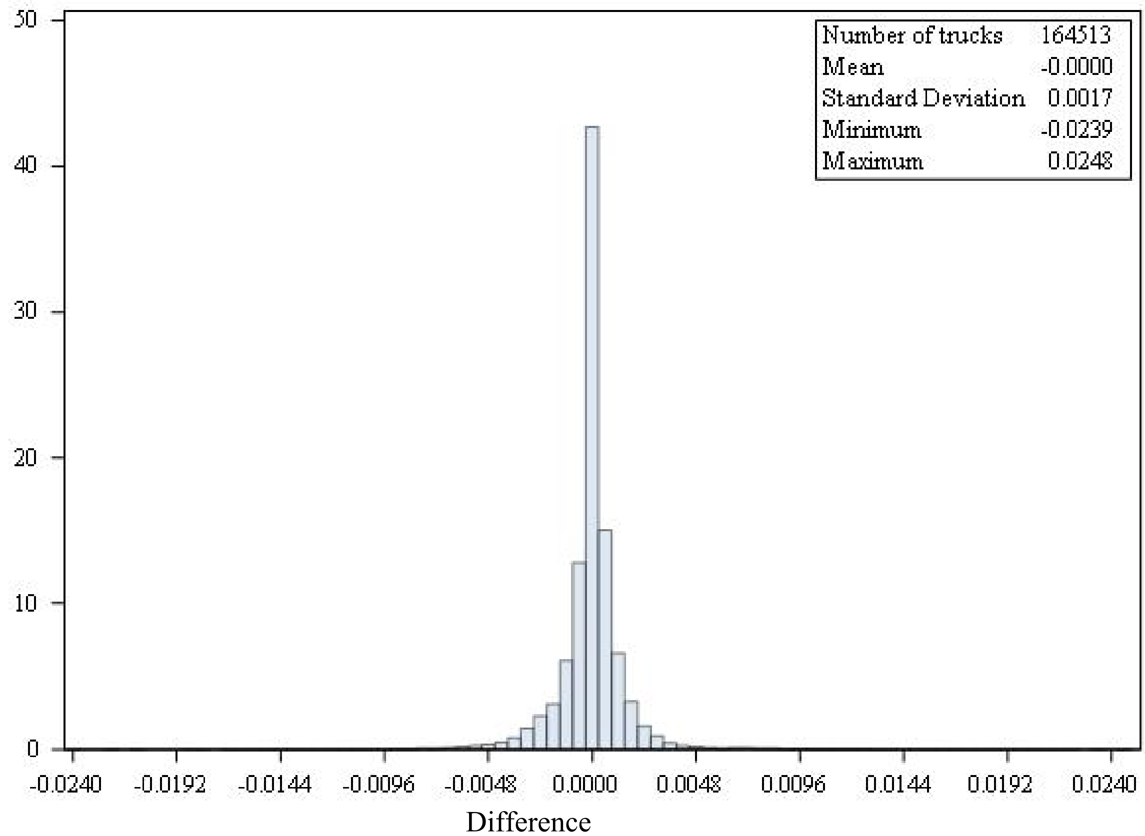

Figure OA5.1: Distribution of the difference between K=18 and K=19 of the posterior expected number of accidents at period T+1 obtained from the estimation of the hierarchical model.

Other estimation results

Estimation of Hausman and gamma Dirichlet models with fleets of one truck or more

Table OA5.7: Parameter estimation for the distribution of the number of annual truck accidents for the 1991–1998 period, for fleets with one truck or more, and trucks with two periods or more: Gamma-Dirichlet model and Hausman model

	Gamma-Dirichlet		Hausman model	
Explanatory variable	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.7159*	0.0410	-0.8149*	0.0517
Number of years as a fleet	-0.0470*	0.0028	-0.0456*	0.0029
Sector of activity in 1998				
Other sectors	-0.1593	0.0868	-0.2068	0.0857
General public trucking	0.1479*	0.0244	0.0334	0.0233
Bulk public trucking	Reference group			
Private trucking	0.0950*	0.0198	0.0858*	0.0184
Short-term rental firm	0.4589*	0.0392	0.3944*	0.0326
Size of fleet				
1	-0.0398*	0.0154	-0.0580*	0.0149
2	Reference group			
3	0.1190*	0.0187	0.1253*	0.0185
4 to 5	0.1933*	0.0183	0.1919*	0.0175
6 to 9	0.2836*	0.0185	0.2750*	0.0175
10 to 20	0.3786*	0.0181	0.3517*	0.0168
21 to 50	0.3790*	0.0193	0.3527*	0.0181
More than 50	0.3048*	0.0186	0.3563*	0.0169
Days in previous year	1.8841*	0.0251	1.6236*	0.0248
Violations				
For overload	0.0966*	0.0096	0.1181*	0.0104
For excess size	0.1799	0.0763	0.1519	0.0851
For poorly secured cargo	0.1710*	0.0271	0.2866*	0.0335
Not respect service hours	0.1787*	0.0620	0.2492*	0.0658
No mechanical inspection	0.2191*	0.0262	0.2519*	0.0268
For other reasons	0.1558*	0.0667	0.2467*	0.0737
Type of vehicle use				
Commercial use	-0.2858*	0.0182	-0.1801*	0.0183
Other than bulk goods	-0.1528*	0.0212	-0.0712*	0.0223
Bulk goods	Reference group			
Type of fuel				
Diesel	Reference group			
Gas	-0.3980*	0.0117	-0.4601*	0.0122

Explanatory variable	Gamma-Dirichlet Coefficient	Gamma-Dirichlet Standard error	Hausman model Coefficient	Hausman model Standard error
Other	-0.2689*	0.0687	-0.3402*	0.0724
Number of cylinders				
1 to 5	0.2762*	0.0339	0.2995*	0.0355
6 to 7	0.3388*	0.0112	0.3612*	0.0121
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.0493*	0.0168	-0.2127*	0.0177
2 axles (4,000 kg or more)	-0.0503*	0.0124	-0.2313*	0.0137
3 axles	0.0414*	0.0124	-0.2189*	0.0135
4 axles	-0.0331	0.0163	-0.1841*	0.0181
5 axles	-0.0231	0.0149	-0.2255*	0.0156
6 axles or more	Reference group			
Number of violations				
For speeding	0.1786*	0.0077	0.2561*	0.0092
Suspended license	0.3717*	0.0327	0.4327*	0.0335
For running a red light	0.2679*	0.0198	0.4343*	0.0213
For ignoring a stop sign	0.4220*	0.0219	0.4788*	0.0222
Not wearing a seat belt	0.1991*	0.0249	0.2391*	0.0256
Observation period				
1991	-0.0252	0.0223	0.0267	0.0560
1992	-0.0282	0.0201	-0.0090	0.0242
1993	0.0183	0.0185	-0.0720*	0.0221
1994	0.1258*	0.0169	-0.0063	0.0200
1995	0.1683*	0.0156	0.0099	0.0175
1996	0.0903*	0.0147	-0.0435	0.0171
1997	-0.0885*	0.0145	-0.1530*	0.0153
1998	Reference group			
$\hat{v}_0$	2.0162*	0.0163		
$\hat{\kappa}_0$	1.4725*	0.0560		
$\hat{\rho}_0$	0.6907*	0.0203		
$\hat{a}$			21.0734*	0.3700
$\hat{b}$			1.6563*	0.0272
Number of observations	678,331		678,331	
Number of trucks	164,513		164,513	
Number of fleets	62,171		62,171	
Log likelihood	-270,956		-269,077	
$\hat{v} = \exp(\hat{v}_0)$	7.5097*			
$\hat{\kappa} = \exp(\hat{\kappa}_0)$	4.3601*			

Explanatory variable	Gamma-Dirichlet Coefficient	Gamma-Dirichlet Standard error	Hausman model Coefficient	Hausman model Standard error
$\hat{\rho} = \exp(\hat{\rho}_0)$	1.9951*			

* Significant at 1%.

Table OA5.8: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991–1996 and 1991–1995 periods, for fleets with one truck or more, and trucks with two periods or more: hierarchical random-effects model with $K = 19$

Explanatory variable	Hierarchical model 91–96 Coefficient	Hierarchical model 91–96 Standard error	Hierarchical model 91–95 Coefficient	Hierarchical model 91–95 Standard error
Constant	-3.2661*	0.0471	-3.1613*	0.0530
Number of years as a fleet	-0.0566*	0.0040	-0.0599*	0.0052
Sector of activity in 1998				
Other sectors	-0.1040	0.1092	-0.1146	0.1179
General public trucking	0.1542*	0.0315	0.1553*	0.0337
Bulk public trucking		Reference group		
Private trucking	0.0745*	0.0236	0.0737*	0.0254
Short-term rental firm	0.4614*	0.0574	0.4335*	0.0564
Size of fleet				
1	-0.0619*	0.0167	-0.0718*	0.0183
2		Reference group		
3	0.1199*	0.0213	0.0985*	0.0235
4 to 5	0.1600*	0.0213	0.1543*	0.0235
6 to 9	0.2187*	0.0224	0.2029*	0.0245
10 to 20	0.2898*	0.0232	0.2690*	0.0253
21 to 50	0.2245*	0.0275	0.1971*	0.0304
More than 50	0.2459*	0.0277	0.2276*	0.0328
Days in previous year	1.7001*	0.0290	1.6800*	0.0327
Violations				
Overload	0.1117*	0.0118	0.1068*	0.0133
Excess size	0.1339	0.0884	0.1190	0.1019
Poorly secured cargo	0.2212*	0.0374	0.2424*	0.0408
Not respecting service hours	0.1738	0.0844	0.2130	0.0960
No mechanical inspection	0.2120*	0.0286	0.2469*	0.0309
Other reasons	0.1463	0.0846	0.1229	0.0950
Type of vehicle use				
Commercial use	-0.1948*	0.0219	-0.1800*	0.0239
Other than bulk goods	-0.0668	0.0264	-0.0419	0.0286
Bulk goods		Reference group		
Type of fuel				
Diesel		Reference group		
Gas	-0.4275*	0.0131	-0.4236*	0.0140

	Hierarchical model 91–96		Hierarchical model 91–95	
Explanatory variable	Coefficient	Standard error	Coefficient	Standard error
Other	-0.3533*	0.0840	-0.3055*	0.0900
Number of cylinders				
1 to 5	0.2457*	0.0397	0.2285*	0.0440
6 to 7	0.3261*	0.0129	0.3170*	0.0139
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3231*	0.0210	-0.3168*	0.0139
2 axles (more than 4,000 kg)	-0.3538*	0.0166	-0.3520*	0.0182
3 axles	-0.2885*	0.0163	-0.2928*	0.0178
4 axles	-0.1903*	0.0211	-0.2058*	0.0231
5 axles	-0.2348*	0.0189	-0.2352*	0.0208
6 axles or more	Reference group			
Number of violations				
For speeding	0.2324*	0.0115	0.2324*	0.0133
Suspended license	0.4121*	0.0388	0.4313*	0.0464
For running a red light	0.3644*	0.0239	0.3570*	0.0267
For ignoring a stop sign	0.4284*	0.0251	0.4329*	0.0281
Not wearing a seat belt	0.2130*	0.0287	0.2160*	0.0323
Observation period				
1991	0.0284	0.0223	-0.0399	0.0230
1992	0.0038	0.0193	-0.0619*	0.0192
1993	-0.0520*	0.0171	-0.1154*	0.0164
1994	0.0236	0.0151	-0.0363	0.0142
1995	0.0526*	0.0137	Reference group	
1996	Reference group			
$\hat{\delta}$	0.6998*	0.0383	0.6870*	0.0445
$\widehat{\beta c}_0$	2.1509*	0.0691	2.1242*	0.0764
$\hat{\delta}^*$	2.8228*	0.3493	3.1708*	0.4844
$\hat{\beta}^*$	2.7770*	0.3676	3.1436*	0.5152
Number of observations	491,792		397,098	
Number of trucks	132,868		115,280	
Number of fleets	53,088		48,194	
Log likelihood	-196,608		-160,646	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.6826		0.6799	

* Significant at 1%.

Different values of *K*

Table OA5.8 presents the results of the hierarchical random model with *K*=8 and with *K*=10. Table OA5.9 gives the results for *K*=13 and *K*=15. The log likelihood is -266,017 at *K*=8 and increases to -265,790 at *K*=10, a value very close to the log likelihood values at *K*=13 (-265,574) and at *K*=15 (-265,478). We do not see important differences at *K*=8 when compared to *K* values equal to or greater than 10. The coefficient estimates do not vary very much with the exception of the random effects parameters. However, the posterior expected number of accidents at *T*+1 differs between different *K* values, contrarily to the comparison presented between *K*=18 and *K*=19 in Section 6, where it is stable.

Table OA5.9: Parameters estimation of the distribution of the number of annual truck accidents for the 1991-1998 period, for fleets with one truck or more, and trucks with two periods or more: hierarchical random effects models with *K*=8 and *K*=10.

Explanatory variables	*K*=8		*K*=10	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.1897*	0.0386	-3.1845*	0.0389
Number of years as a fleet on December 31st	-0.0493*	0.0025	-0.0494	0.0026
Sector of activity in 1998				
Other sectors	-0.2241	0.0963	-0.2249	0.0939
General public trucking	0.0890*	0.0250	0.0899*	0.0253
Bulk public trucking	Reference group			
Private trucking	0.0491	0.0195	0.0465	0.0197
Short-term rental firm	0.3743*	0.0433	0.3727*	0.0430
Size of fleet				
1	-0.0503*	0.0144	-0.0505*	0.0145
2	Reference group			
3	0.1165*	0.0183	0.1174*	0.0184
4 to 5	0.1664*	0.0179	0.1687*	0.0181
6 to 9	0.2182*	0.0185	0.2227*	0.0188
10 to 20	0.2415*	0.0187	0.2516*	0.0192
21 to 50	0.1735*	0.0215	0.1908*	0.0220
More than 50	0.1486*	0.0198	0.1998*	0.0202
Days in previous year	1.6432*	0.0245	1.6453*	0.0244
Violations				
Overload	0.1050*	0.0103	0.1029*	0.0103
Excessive size	0.1568	0.0768	0.1576	0.0766

Explanatory variables	K=8		K=10	
	Coefficient	Standard error	Coefficient	Standard error
Poorly secured cargo	0.2627*	0.0337	0.2569*	0.0335
Not respecting service hours	0.2145*	0.0662	0.2100*	0.0660
No mechanical inspection	0.2289*	0.0261	0.2247*	0.0261
Other reasons	0.2230*	0.0702	0.2216*	0.0700
Type of vehicle use				
Commercial use	-0.1957*	0.0188	-0.1949*	0.0189
Other than bulk goods	-0.0729*	0.0226	-0.0760*	0.0227
Bulk goods	Reference group			
Type of fuel				
Diesel	Reference group			
Gas	-0.4460*	0.0121	-0.4420*	0.0121
Other	-0.3208*	0.0729	-0.3217*	0.0729
Number of cylinders				
1 to 5	0.2442*	0.0351	0.2407*	0.0351
6 to 7	0.3381*	0.0118	0.3353*	0.0118
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3243*	0.0184	-0.3324*	0.0185
2 axles (4,000 kg and more)	-0.3569*	0.0144	-0.3649*	0.0144
3 axles	-0.3084*	0.0139	-0.3081*	0.0139
4 axles	-0.2185*	0.0185	-0.2177*	0.0185
5 axles	-0.2590*	0.0159	-0.2576*	0.0159
6 axles or more	Reference group			
Number of violations				
Speeding	0.2341*	0.0093	0.2293*	0.0093
Suspended license	0.4018*	0.0332	0.3982*	0.0332
Running a red light	0.4003*	0.0209	0.3940*	0.0209
Ignoring a stop sign	0.4313*	0.0221	0.4259*	0.0220
Not wearing a seat belt	0.2216*	0.0249	0.2182*	0.0249
Observation period				
1991	0.0021	0.0205	0.0028	0.0206
1992	-0.0273	0.0185	-0.0274	0.0186
1993	-0.0902*	0.0177	-0.0900*	0.0173
1994	-0.0216	0.0159	-0.0213	0.0159
1995	0.0007	0.0148	0.0008	0.0148
1996	-0.0467*	0.0141	-0.0471*	0.0141
1997	-0.1549*	0.0140	-0.1554*	0.0140
1998	Reference group			
$\hat{\delta}$	0.4131*	0.0180	0.4593*	0.0204
$\widehat{\beta c_0}$	1.1288*	0.0302	1.3828*	0.0371
$\hat{\delta}^*$	3.6213*	0.3504	3.5225*	0.3542
$\hat{\beta}^*$	5.4061*	0.5508	4.6294*	0.4914

Explanatory variables	K=8		K=10	
	Coefficient	Standard error	Coefficient	Standard error
Number of observations	678,331		678,331	
Number of trucks	164,513		164,513	
Number of fleets	62,171		62,171	
Log likelihood	-266,017		-265,790	
$\widehat{\beta c} = \dfrac{\widehat{\beta c_0}}{1 + \widehat{\beta c_0}}$	0.5303		0.5803	

* Significant at 1%.

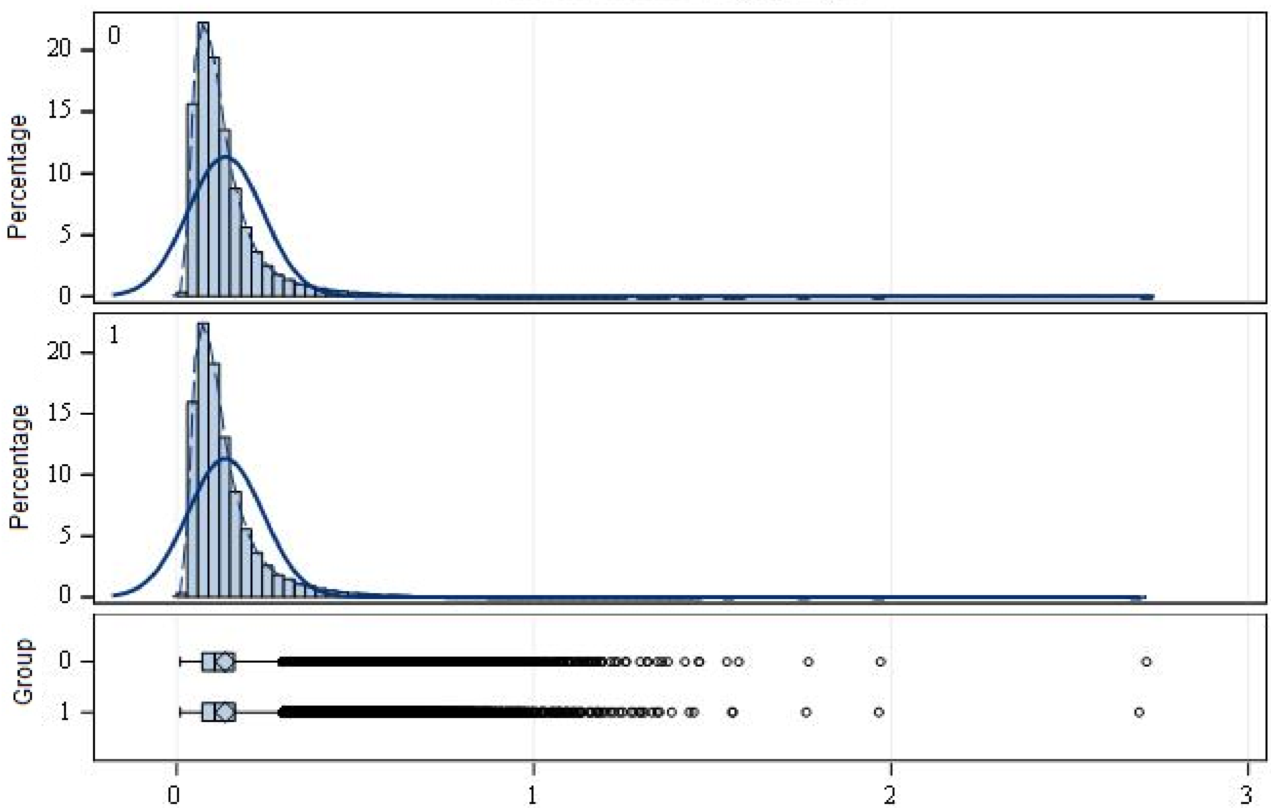

Figure OA5.2: Posterior expected number of accidents at period *T*+1 obtained from the estimation of hierarchical model at *K*=8 (at top) and at *K*=10 (at bottom)

We can see from Figure OA5.2 that the distribution of the posterior expected number of accidents at period *T*+1 with *K*=10 (at bottom) and with *K*=8 (at top). The mean of the difference is statistically different from zero at 1%, as shown in Figure OA5.3: the *t*-test value is equal 7.93 and the p-value is less than 0.0001.

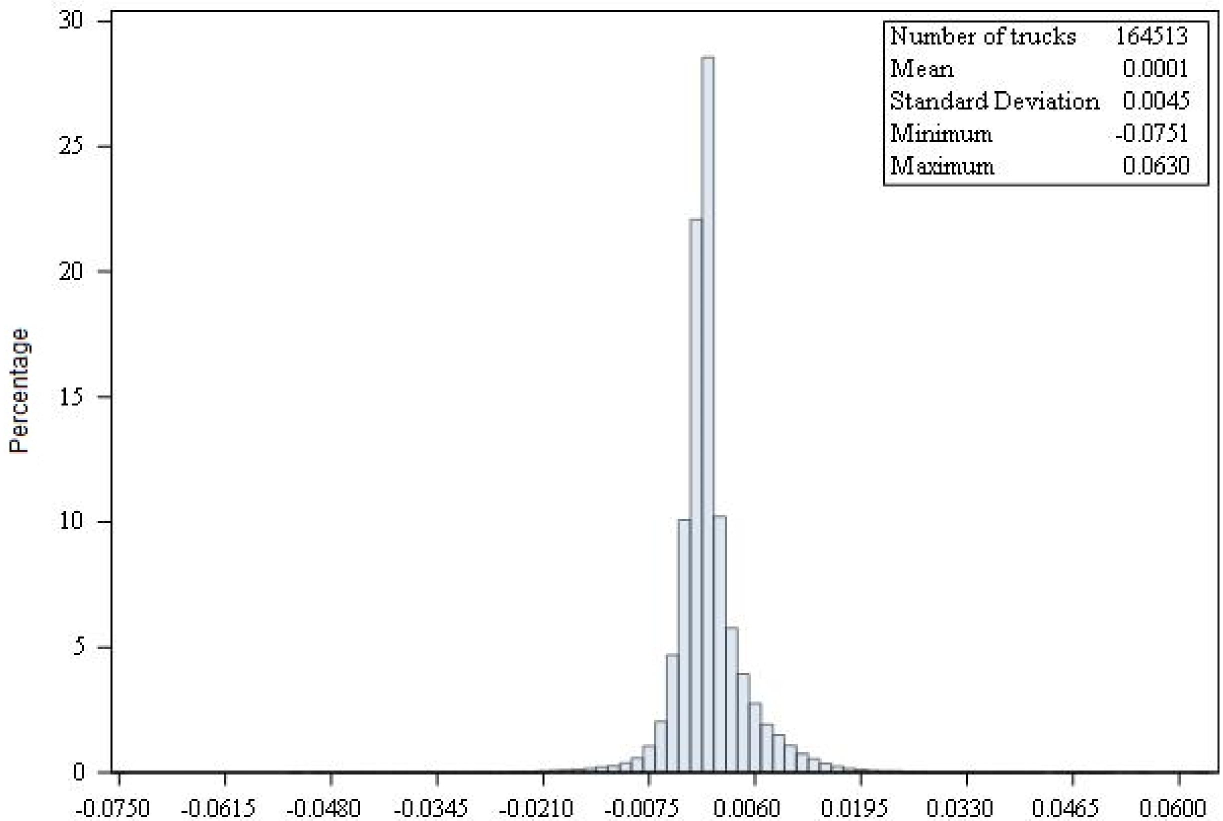

Figure OA5.3: Distribution of the difference between K=8 and K=10 of the posterior expected number of accidents at period T+1 obtained from the estimation of the hierarchical model.

Table OA5.10: Parameters estimation of the distribution of the number of annual truck accidents for the 1991-1998 period, for fleets with one truck or more, and trucks with two periods or more: hierarchical random effects models with K = 13 and K = 15.

Explanatory variables	K=13		K=15	
	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.1730*	0.0394	-3.1775*	0.0398
Number of years as a fleet on December 31st	-0.0500*	0.0026	-0.0501*	0.0027
Sector of activity in 1998				
Other sectors	-0.2122	0.0939	-0.2043	0.0956
General public trucking	0.1032*	0.0262	0.0957*	0.0267
Bulk public trucking	Reference group			
Private trucking	0.0448	0.0200	0.0473	0.0202
Short-term rental firm	0.4175*	0.0428	0.4530*	0.0453
Size of fleet				
1	-0.0512*	0.0146	-0.0519*	0.0146
2	Reference group			
3	0.1182*	0.0185	0.1191*	0.0185
4 to 5	0.1709*	0.0183	0.1732*	0.0183
6 to 9	0.2271*	0.0191	0.2313*	0.0192
10 to 20	0.2610*	0.0197	0.2695*	0.0199
21 to 50	0.2071*	0.0227	0.2200*	0.0232
More than 50	0.2171*	0.0219	0.2224*	0.0238
Days in previous year	1.6456*	0.0244	1.6474*	0.0245
Violations				
Overload	0.1004*	0.0103	0.1000*	0.0103
Excessive size	0.1578	0.0764	0.1583	0.0764
Poorly secured cargo	0.2499*	0.0333	0.2473*	0.0332
Not respecting service hours	0.2037*	0.0657	0.2023*	0.0655
No mechanical inspection	0.2213*	0.0260	0.2197*	0.0259
Other reasons	0.2196*	0.0698	0.2189*	0.0697
Type of vehicle use				
Commercial use	-0.1928*	0.0190	-0.1920*	0.0190
Other than bulk goods	-0.0742*	0.0228	-0.0769*	0.0229
Bulk goods	Reference group			
Type of fuel				
Diesel	Reference group			
Gas	-0.4384*	0.0121	-0.4358*	0.0120
Other	-0.3278*	0.0731	-0.3256*	0.0737
Number of cylinders				
1 to 5	0.2337*	0.0349	0.2275*	0.0347
6 to 7	0.3323*	0.0117	0.3314*	0.0117
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3423*	0.0185	-0.3426*	0.0184

Explanatory variables	K=13		K=15	
	Coefficient	Standard error	Coefficient	Standard error
2 axles (4,000 kg and more)	-0.3794*	0.0145	-0.3802*	0.0145
3 axles	-0.3092*	0.0140	-0.3042*	0.0141
4 axles	-0.2173*	0.0186	-0.2125*	0.0187
5 axles	-0.2603*	0.0161	-0.2552*	0.0162
6 axles or more	Reference group			
Number of violations				
Speeding	0.2246*	0.0093	0.2227*	0.0093
Suspended license	0.3965*	0.0331	0.3967*	0.0331
Running a red light	0.3893*	0.0208	0.3867*	0.0207
Ignoring a stop sign	0.4224*	0.0220	0.4203*	0.0219
Not wearing a seat belt	0.2169*	0.0248	0.2158*	0.0248
Observation period				
1991	-0.0018	0.0209	-0.0012	0.0211
1992	-0.0312	0.0188	-0.0308	0.0190
1993	-0.0927*	0.0175	-0.0918*	0.0176
1994	-0.0229	0.0160	0.0221	0.0161
1995	-0.0001	0.0149	-0.0001	0.0149
1996	-0.0478*	0.0141	-0.0480*	0.0142
1997	-0.1558*	0.0140	-0.1559*	0.0140
1998	Reference group			
$\hat{\delta}$	0.5498*	0.0245	0.6154*	0.0281
$\widehat{\beta c}_0$	1.6839*	0.0453	1.8587*	0.0504
$\hat{\delta}^*$	3.2283*	0.3283	2.9859*	0.3053
$\hat{\beta}^*$	3.6640*	0.3970	3.1536*	0.3426
Number of observations	678,331		678,331	
Number of trucks	164,513		164,513	
Number of fleets	62,171		62,171	
Log likelihood	-265,574		-265,478	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.6274		0.6502	

* Significant at 1%.

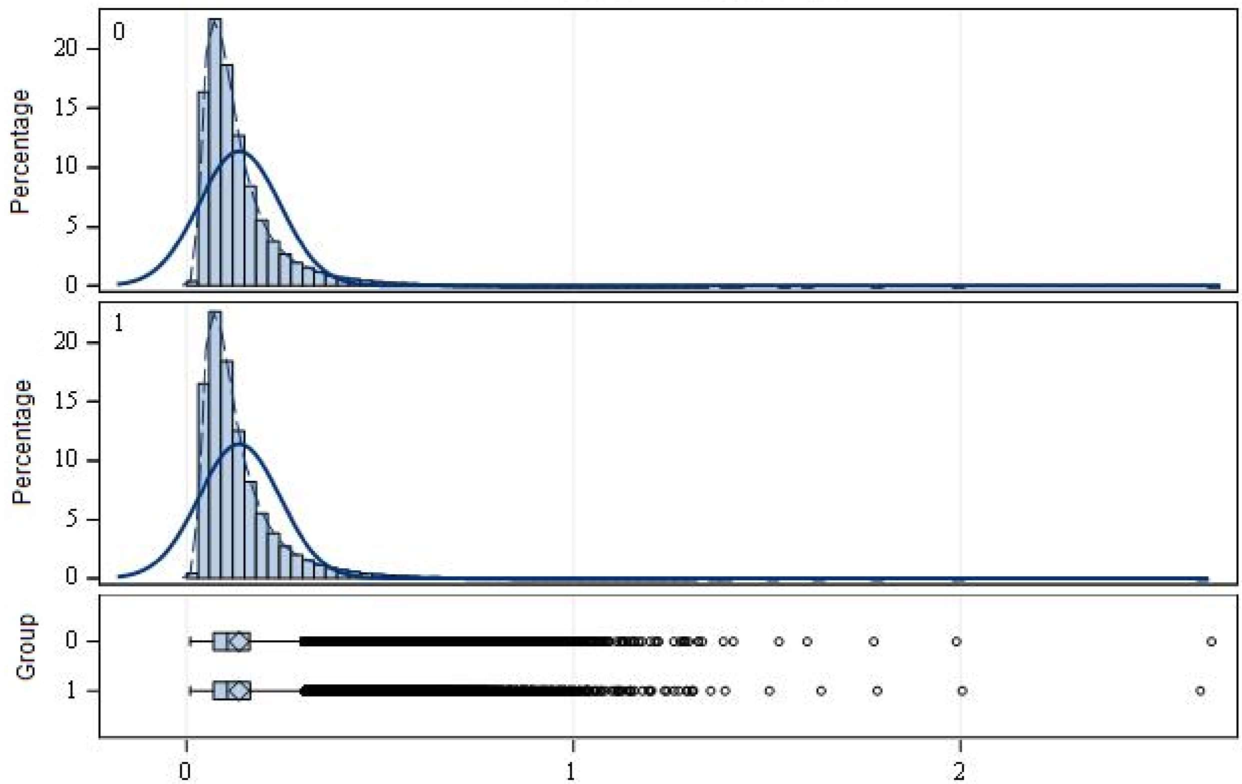

Figure OA5.4: Posterior expected number of accidents at period T+1 obtained from the estimation of hierarchical model at K=13 (at top) and at K=15 (at bottom)

We can see from Figure OA5.4 the distribution of the posterior expected number of accidents at period T+1 with K=13 (at bottom) and with K=15 (at top). Figure OA5.5 presents the distribution of the difference used for the paired t-test. The mean of the difference is statistically different from zero at 1%: the t-test value is equal 6.64 and the p-value is less than 0.0001.

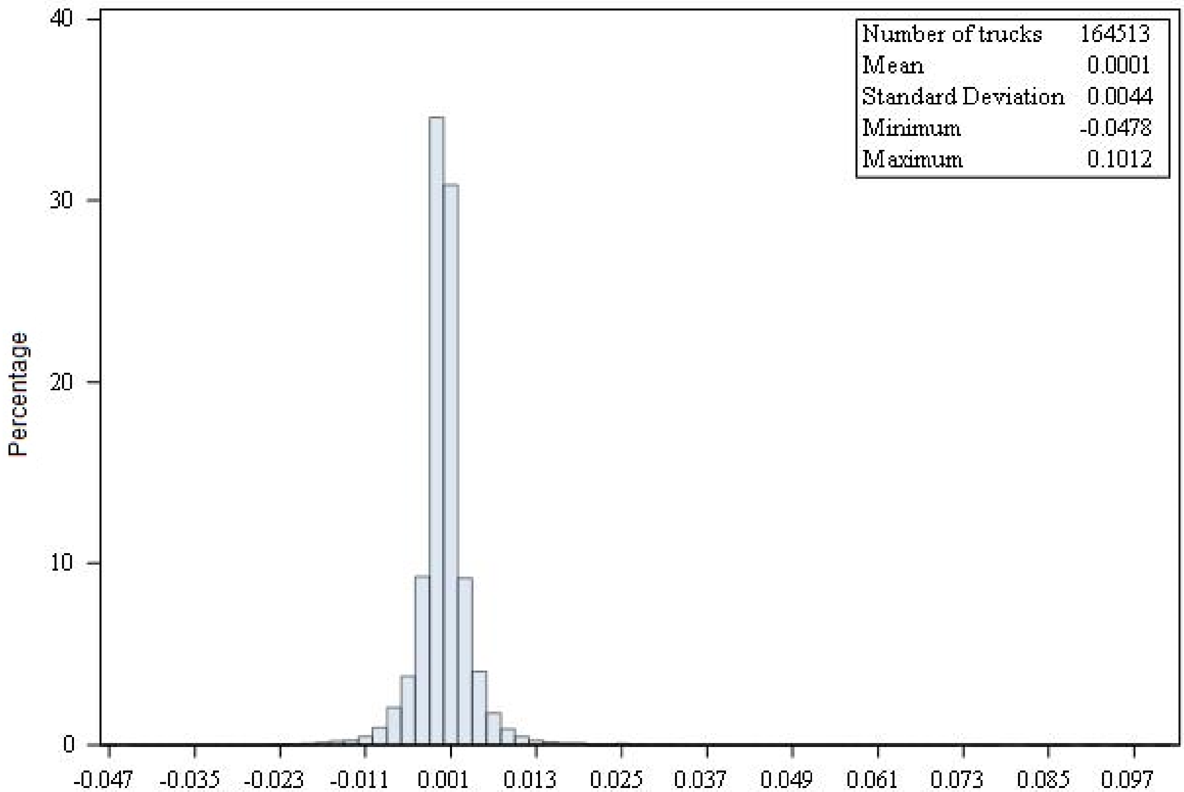

Figure OA5.5: Distribution of the difference between *K*=13 and *K*=15 of the posterior expected number of accidents at period *T*+1 obtained from the estimation of hierarchical model.

Table OA5.11: Estimation of the parameters of the distribution of the number of annual truck accidents for the 1991–1998 period, for fleets with two trucks or more and trucks with two periods or more: Hausman, gamma-Dirichlet and hierarchical model (*K*=13).

Explanatory variables	Hausman model		Gamma-Dirichlet model		Hierarchical	
	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 sectors	-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*	0.0196
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						
Overload	0.1216*	0.0117	0.0966*	0.0115	0.0809*	0.0119
Excessive size	0.1456***	0.0883	0.1480***	0.0860	0.1448	0.0884
Poorly secured cargo	0.2522*	0.0363	0.2054*	0.0354	0.1826*	0.0365
Not respecting 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
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	0.3541*	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	
	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						
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
Running a red light	0.3838*	0.0247	0.3094*	0.0239	0.2815*	0.0246
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	0.0233	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		-193,036	
$\widehat{\beta c} = \frac{\widehat{\beta c_0}}{1 + \widehat{\beta c_0}}$					0.7531	

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

Table OA5.12 Fit statistics of the three models in Table OA5.8 for fleets of two trucks or more

Statistics	Hausman model	Gamma-Dirichlet model	Hierarchical 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: The likelihood ratio test value of 8,160 is largely superior to the critical value of 6.63 at 1% when comparing the gamma-Dirichlet model to the hierarchical model. The likelihood ratio test value of 98 is superior to the same critical value when comparing the Hausman model to the gamma-Dirichlet model. The likelihood ratio test value of 8,254 is superior to the critical value of 9.21 at 1% when comparing the Hausman model to the hierarchical model with *K*=13.

Table OA5.13: Parameters estimation of the distribution of the number of annual truck accidents for the 1991-1997 and 1991-1996 periods, for fleets with one truck or more and trucks with two periods or more: hierarchical random effects models with $K = 19$.

	Hierarchical model 91-97		Hierarchical model 91-96	
Explanatory variable	Coefficient	Standard error	Coefficient	Standard error
Constant	-3.3133*	0.0433	-3.2661*	0.0471
Number of years as a fleet	-0.0553*	0.0033	-0.0566*	0.0040
Sector of activity in 1998				
Other sectors	-0.1783	0.1031	-0.1040	0.1092
General public trucking	0.1326*	0.0287	0.1542*	0.0315
Bulk public trucking		Reference group		
Private trucking	0.0665*	0.0218	0.0745*	0.0236
Short-term rental firm	0.4926*	0.0517	0.4614*	0.0574
Size of fleet				
1	-0.0633*	0.0156	-0.0619*	0.0167
2		Reference group		
3	0.1161*	0.0199	0.1199*	0.0213
Sizes 4 to 5	0.1767*	0.0198	0.1600*	0.0213
Size 6 to 9	0.2240*	0.0208	0.2187*	0.0224
Sizes 10 to 20	0.2742*	0.0217	0.2898*	0.0232
Sizes 21 to 50	0.2300*	0.0257	0.2245*	0.0275
Sizes > 50	0.2470*	0.0258	0.2459*	0.0277
Days in previous year	1.6535*	0.0262	1.7001*	0.0290
Violations				
Overload	0.1005*	0.0107	0.1117*	0.0118
Excessive size	0.1534	0.0790	0.1339	0.0884
Poorly secured cargo	0.2282*	0.0349	0.2212*	0.0374
Not respecting service hours	0.2002*	0.0751	0.1738	0.0844
No mechanical inspection	0.2096*	0.0271	0.2120*	0.0286
Other reasons	0.2167	0.0752	0.1463	0.0846
Type of vehicle use				
Commercial use	-0.1955*	0.0203	-0.1948*	0.0219
Other than bulk goods	-0.0767*	0.0244	-0.0668	0.0264
Bulk goods		Reference group		
Type of fuel				
Diesel		Reference group		
Gas	-0.4223*	0.0125	-0.4275*	0.0131
Other	-0.3852*	0.0800	-0.3533*	0.0840

Explanatory variable	Hierarchical model 91-97 Coefficient	Standard error	Hierarchical model 91-96 Coefficient	Standard error
Number of cylinders				
1 to 5	0.2327*	0.0369	0.2457*	0.0397
6 to 7	0.3291*	0.0122	0.3261*	0.0129
8 or more than 10	Reference group			
Number of axles				
2 axles (3,000 to 4,000 kg)	-0.3300*	0.0195	-0.3231*	0.0210
2 axles (4,000 kg and more)	-0.3693*	0.0154	-0.3538*	0.0166
3 axles	-0.2944*	0.0151	-0.2885*	0.0163
4 axles	-0.2034*	0.0198	-0.1903*	0.0212
5 axles	-0.2454*	0.0174	-0.2348*	0.0189
6 axles or more	Reference group			
Number of violations				
Speeding	0.2245*	0.0101	0.2324*	0.0115
Suspended license	0.3863*	0.0349	0.4121*	0.0388
Running a red light	0.3636*	0.0223	0.3644*	0.0239
Ignoring a stop sign	0.4155*	0.0235	0.4284*	0.0251
Not wearing a seat belt	0.2262*	0.0257	0.2130*	0.0287
Observation period				
1991	0.1292*	0.0223	0.0284	0.0223
1992	0.1035*	0.0198	0.0038	0.0193
1993	0.0469*	0.0180	-0.0520*	0.0171
1994	0.1213*	0.0162	0.0236	0.0151
1995	0.1481*	0.0149	0.0526*	0.0137
1996	0.1038*	0.0141	Reference group	
1997	Reference group			
$\hat{\delta}$	0.7139*	0.0359	0.6998*	0.0383
$\widehat{\beta c}_0$	2.1638*	0.0643	2.1509*	0.0691
$\hat{\delta}^*$	2.6544*	0.2909	2.8228*	0.3493
$\hat{\beta}^*$	2.5672*	0.3002	2.7770*	0.3676
Number of observations	585,398		491,792	
Number of trucks	149,231		132,868	
Number of fleets	57,741		53,088	
Log likelihood	-230,373		-196,608	
$\widehat{\beta c} = \frac{\widehat{\beta c}_0}{1 + \widehat{\beta c}_0}$	0.6839		0.6826	

* Significant at 1%.

Table OA5.14: *t*-test of the posterior expected number of accidents at period *T*+1 from estimation of hierarchical model 91-97 and the observed numbers of accidents in 1998 for all fleets.

	Hierarchical model 91-97			Data 1998			*t*-test	
	N trucks	Mean	Std	N trucks	Mean	Std	*t*-value	*p*-value
All fleets	149,231	0.1373	0.1045	77,651	0.1418	0.4109	-2.96	0.0031
Size 1	36,198	0.0990	0.0689	25,230	0.1085	0.3581	-4.16	<0.0001
Size 2	21,082	0.1107	0.0811	9,726	0.1092	0.3492	0.41	0.6795
Size 3	11,793	0.1192	0.0878	5,922	0.1348	0.3981	-2.97	0.0030
Sizes 4 to 5	14,287	0.1337	0.0967	7,321	0.1488	0.4521	-2.99	0.0028
Size 6 to 9	14,273	0.1526	0.1136	6,819	0.1684	0.4401	-2.91	0.0036
Sizes 10 to 20	15,564	0.1731	0.1256	7,754	0.1893	0.4751	-2.96	0.0030
Sizes 21 to 50	13,490	0.1812	0.1213	6,371	0.1874	0.4742	-1.03	0.3011
Sizes > 50	22,544	0.1749	0.1181	8,508	0.1776	0.4652	-0.52	0.6024

Mean 91-97: Posterior expected number of accidents in year 1998 from estimations in Table OA5.10 (hierarchical model 91-97)
Data 1998: Observed mean of accidents in 1998