# Nonlinear Panel Data Models with Dynamic Heterogeneity


Christian Gourieroux[1]

Joanna Jasiak[2]



This version: November 23, 1998

(First draft: October 12, 1998)



[1]CREST and CEPREMAP , *e-mail*: `gouriero@ensae.fr`.

[2]York University, *e-mail*: `jasiakj@yorku.ca`

**Abstract**

Nonlinear Panel Data Models with Dynamic Heterogeneity.

We introduce nonlinear panel data models for individual risk management and control. Our approach extends traditional models by allowing for nonlinearities in input variables and unobserved individual heterogeneity with possible temporal dependence. This permits us to develop nonlinear models for the assessment of individual risks in the presence of moral hazard and adverse selection. In practice, these phenomena are confronted by banks and insurance companies in their basic activities of fund supply and loss coverage. Although moral hazard and adverse selection are not perfectly controllable, their effects can be alleviated by adjustments to the pure premiums and an updating scheme based on the approximated risk.



Keywords: Panel, Factor Models, Credibility Theory, Insurance, Poisson-Gamma Model

JEL : C33, G22.

**Résumé**

Modèles de panel non linéaires avec hétérogénéité dynamique.

Nous considérons des modèles de panel non linéaires, où les facteurs d' hétérogénéité individuelle non observable peuvent être corrélés dans le temps. Ces modèles sont particulièrement utiles pour l'analyse des risques individuels en crédit ou en assurance, car ils permettent de tenir compte des phénomènes de hazard moral et de modifier en conséquence les calculs de primes et leurs schémas de mise à jour.

Mot clés: panel, modèles à facteurs, théorie de la crédibilité, assurance, modèle Poisson-Gamma.

JEL : C33, G22.

# 1 Introduction

An important area of application of nonlinear panel models involves the assessment of individual risks [1]. Recently, a growing concern of financial institutions is the increasing liability and falling insurance efficiency attributable in part to moral hazard and adverse selection [2]. These phenomena are difficult to control and obscure appropriate risk evaluation. As they arise because of unsufficient information available to financial institutions, the ultimate remedy consists in acquiring more knowledge about borrowers or policyholders. The strategy of banks and insurance companies is focused on learning from past experiences. Indeed, various financial and insurance products have been developed which maintain the customers wealth but allow the companies to accumulate data on the history of behavior of borrowers or policyholders. Revolving credit, where the borrower is offered a permanent varying debt, is an example of such product, since it eliminates the cost of multiple individual decisions on borrowing and new agreements on repayment. Similarly, motor vehicle insurers reward low risk customers by updating their policy premiums and granting them bonuses. These examples show that banks and insurers seek to create an environment favoring long term commitments of individuals.

It is commonly assumed that the distribution of individual risks depends on three types of factors:

i) Observed exogenous characteristics of the individual or of the contract selected by that person reveal her/his risk category.

The standard practice of credit institutions or insurance companies consists in using these characteristics to define a priori classes of homogenous risks. Within these classes any residual effect of the exogenous variables can be neglected.

ii) Updated approximations of unobserved individual factors are necessary to alleviate the asymmetry of information in favour of the individual.

This unobserved component is more accurately approximated when the observed risk history increases. The updating of this approximation underlies the current bonus-malus scheme introduced, for instance, in car insurance.

iii) Individual effort to prevent losses is unobserved and features temporal dependence.

The individual may adjust his/her effort for loss prevention according to his/her experience with past claims, the price of the contract and awareness of future consequences of an incident. The effort variable determines the moral hazard and is modeled by a dynamic unobserved factor.

The total amount of claims, confronted by companies in a fixed period of time is an outcome of several variables of different types. The dichotomous qualitative variables indicate the occurrence

[1] See selected applications in Matyas, Sevestre (1996)

[2] See, Geneva Papers on Risk and Insurance, No.77, October 1995, or Loubergé (1998)

of claims, the discrete variables measure the number of claims, and finally the quantitative variables with a point mass at zero, capture the total cost of claims. Due to this variety of components, *nonlinear models* are necessary to represent the magnitude of covered losses incurred in a given period.

In this paper we introduce nonlinear panel models with dynamic heterogeneity for individual risk analysis. Our approach allows for unobserved individual factors and permits us to evaluate the comparative effects of adverse selection and moral hazard. We emphasize the consequences of model specification for insurance pricing in terms of pure premiums, the premium updating and the Value at Risk.

The paper is organized as follows. In section 2, we present the linear credibility theory introduced into the actuarial literature by Bailey (1945), (1950), Buhlmann (1967), (1979), and Buhlmann and Straub (1970). We give the explicit form of the associated linear panel model and discuss its limitations affecting both the premium updating scheme and the description of individual structural behavior [see, Balestra, Nerlove (1996), Balestra (1996)]. Significant improvements are obtained by extending the model and introducing dynamic heterogeneity. In particular, we illustrate the advantage of including an unobserved, autoregressive factor of order one. In section 3, we propose nonlinear panel models with dynamic heterogeneity. In this approach, the unobserved heterogeneity is not uniquely defined. We use the example of car insurance, where the risk is measured by a bivariate endogenous variable consisting of the number of claims and their total cost to explain how the initial heterogeneity factors can be transformed into gaussian heterogeneity factors.

This paper presents an analysis performed within a class of homogenous risks, eliminating the presence of explanatory variables. The partition of risks into homogenous classes is out of the scope of this paper. We refer interested readers to Gourieroux (1998a).

## 2 Linear Credibility Theory

We consider a set of individuals $i = 1, ..., n$. For each period $t = 1, ..., T$, the risk associated with individual $i$ is measured by a unidimensional quantitative variable $Y_{i,t}$. In applications to loan and insurance management, the individuals $i$ are assumed to belong to the same homogenous class of risk. This implies in particular that their contracts are of the same type, start at approximately the same date, and the individuals are of the same age when the contract becomes effective. The variable $t$ is the index measuring the time since the beginning of the contract and $Y_{i,t}$ denotes the total cost of the claims for period $t$. The linear credibility theory assumes implicitely that these variables admit continuous distributions on the real line (in some cases an even stronger assumption of gaussian distributions is imposed). This assumption is quite unrealistic considering the nature of dependent variables in the credit and insurance analysis. It is introduced in the next subsection for simplicity of exposition, and is later relaxed in the context of complex nonlinear panel models.

## 2.1 The basic model

In general, the basic model is presented in the literature in a semi-parametric framework [ see, Buhlmann (1967), (1969), Buhlmann, Straub (1970) ]. Nevertheless, its implementation requires the full specification of the distribution, especially when the Value at Risk needs to be determined. A possible choice is the gaussian distribution, which underlies the model discussed below. The model is a simple linear panel model with individual heterogeneity and without explanatory variables. The endogenous variables are such that:

$$Y_{i,t} = m + u_i + v_{i,t}, \quad i = 1, ..., n, \; t = 1, ..., T, ..., \tag{2.1}$$

where:

i ) The two error terms $u_i, \; i = 1, ..., n, \; v_{i,t}, \; i = 1, ..., n, \; t = 1, ..., T$ are independent;

ii) The term $u_i$ representing heterogeneity is gaussian, has zero mean and a fixed variance $\eta^2$;

iii) The error terms $v_{i,t}$ are gaussian with zero mean, and a fixed variance $\sigma^2$.

The main objective of risk analysis is prediction of future risks. It becomes feasible whenever the risks are observed over some period of time, say from time 0 until time $T$, and no attrition is involved. More precisely, for a given horizon $h$ we evaluate the expected total cost $E_T(Y_{i,T+h})$, also called the pure premium, along with the variability of this cost $V_T(Y_{i,T+h})$, and the corresponding quantiles, called the Values at Risk and defined by: $P_T[Y_{i,T+h} > VaR(T, h, \alpha)] = \alpha$. The indexed variables $E_T, V_T, P_T$ indicate that computations are performed conditional on the information available at time $T$.

Under the basic specification the conditional distribution of $(Y_{i,T+h})$ given $I_T = [Y_{i,t}, \; i = 1, ..., n, \; t = 1, ..., T]$ is gaussian. Its mean is:

$$\begin{aligned} E[Y_{i,T+h}|\, I_T] &= E[Y_{i,T+h}|\, Y_{i,1}, ..., Y_{i,T}] \\ &= m + E[u_i|\, Y_{i,1}, ..., Y_{i,T}] \\ &= m + E[u_i|\, \overline{Y}_{i,T}] \qquad (2.2) \\ \text{(where:} \quad & \overline{Y}_{i,T} = \frac{1}{T}\sum_{t=1}^{T} Y_{i,t}) \\ &= m + \frac{Cov(u_i, \overline{Y}_{i,T})}{V(\overline{Y}_{i,T})}(\overline{Y}_{i,T} - m) \end{aligned}$$

$$= \quad m + \frac{\eta^2}{\eta^2 + \sigma^2/T}(\overline{Y}_{i,T} - m). \tag{2.3}$$

The conditional variance is given by:

$$V[Y_{i,T+h} | I_T] = \frac{\sigma^2/T}{\eta^2 + \sigma^2/T}. \tag{2.4}$$

The equation (2.2) defines the pure premium formula, i.e. the basic concept of the credibility theory. The pure premium is a weighted average of the expected cost $m$ for the specific class of risks and the individual empirical average cost $\overline{Y}_{i,T}$:

$$\pi_{i,T}(h) = E(Y_{i,T+h} | I_T) = \frac{\eta^2}{\eta^2 + \sigma^2/T}\overline{Y}_{i,T} + \frac{\sigma^2/T}{\eta^2 + \sigma^2/T}m. \tag{2.5}$$

During the coverage period, data are collected while the observed history is increasing. Simultaneously, the pure premium is being modified:

$$\begin{aligned} \pi_{i,T+1}(h-1) - \pi_{i,T}(h) &= E(Y_{i,T+h} | I_{T+1}) - E(Y_{i,T+h} | I_T) \\ = \quad & \left[\frac{\eta^2}{\eta^2 + \sigma^2/(T+1)} - \frac{\eta^2}{\eta^2 + \sigma^2/T}\right][\bar{Y}_{i,T} - m] \\ & + \frac{\eta^2}{\eta^2 + \sigma^2/(T+1)}[\overline{Y}_{i,T+1} - \overline{Y}_{i,T}]. \end{aligned} \tag{2.6}$$

The modification is determined by two terms. The first one is related to the adjustment of weights assigned to the amount of experience. It indicates that a long contract history is more reliable than a short one. The second component results from the updating of individual average costs.

Recall that this section is based on a restrictive model, which assumes that the individual heterogeneity $u_i$ is independent of the time $t$. This implies several drawbacks:

i) From a structural point of view the basic model cannot account for an individual effort, which is endogenously modified in time. In particular the additional risk due to the moral hazard is neglected and its price is set equal to zero arbitrarily.

ii) The informational advantage of the individual vanishes when a long history is observed. Indeed, when $T$ tends to infinity, we get $\lim_{T\to\infty} \overline{Y}_{i,T} - m = u_i$, and the unobserved heterogeneity becomes known.

iii) Moreover, when $T$ is large, the pure premium becomes $\pi_{i,T}(h) \approx \overline{Y}_{i,T}$. If the price of the contract is based on this pure premium, each individual will asymptotically finance his/her own losses and not those of other insured persons. Asymptotically the insurer creates a balance between the claims and payments of a given individual over various periods, instead of a balance between claims and payments of different individuals in a fixed period. In some sense, the task of insurance company becomes management of spendings and savings of individuals. The insurer is not capable to spread the risk between individuals.

ii) Finally the basic model implies that the conditional distribution of future risks does not depend on the horizon $h$.

## 2.2 Dynamic individual heterogeneity

An immediate extension of the basic models is obtained by introducing a time varying effort variable $u_{i,t}$. To illustrate this technique we represent the dynamics of the effort variable as an autoregressive process of order one. The rationale behind this approach is to incorporate in the model consistent adjustments to individual behavior in order to prevent incidents. The model becomes:

$$Y_{i,t} = m + u_{i,t} + v_{i,t}, \quad i = 1, ..., n, \ t = 1, ..., T, ..., \tag{2.7}$$

where:

i) $u_{i,t} = \phi u_{i,t-1} + \eta\sqrt{1-\phi^2}\epsilon_{i,t}, \ i = 1, ..., n, \ t = 1, ..., T$ with $|\phi| < 1$;

ii) the two error terms $\epsilon_{i,t}$ and $v_{i,t}, \ i = 1, ..., n, \ t = 1, ..., T$ are independent as are the initial values of the individual heterogeneity $u_{i,0}, \ i = 1, ..., n$;

iii) the error terms $\epsilon_{i,t}, \ i = 1, ..., n, \ t = 1, ..., T, ...$ are gaussian with zero mean and unit variance;

iv) the error terms $v_{i,t}, \ i = 1, ..., n, \ t = 1, ..., T, ...$ are gaussian with zero mean and variance $\sigma^2$;

v) the initial heterogeneity factors $u_{i,0}, \ i = 1, ..., n$ are gaussian with zero mean and variance $\eta^2$.

The specification given above differs from the standard panel model with individual effects and autocorrelated errors which is usually written as:

$$Y_{i,t} = m + u_i + v_{i,t},$$

where $v_{i,t}$ is autoregressive of order one. It seems important for the purpose of interpretation to impose the temporal dependence on the $u$ term while keeping the intrinsic error term $v_{i,t}$ unchanged. Indeed, even if the effort component and the intrinsic component cannot be identified in the linear framework, they are easily distinguished in the nonlinear framework considered in section 3. Moreover we introduce in model (2.6) a specific parametrization of the autoregressive process, which ensures that the marginal distribution of $u_{i,t}$ is gaussian, with zero mean and variance $\eta^2$. In particular, the marginal distribution does not depend on the autoregressive coefficient $\phi$. In the limiting case $\phi = 1$, the process $u_{i,t}$ is time independent: $u_{i,t} = u_{i,0}, \ t = 1, ..., T$, and stationary.

The unit root case corresponds to the basic model of subsection 2.1. The parameter $\phi$ can be viewed as a degree of effort updating by the individual: the lower $\phi$, the stronger is the individual effort adjustment.

The serial correlation of effort variables results in significant modifications of prediction formulas. Let us consider the pure premium prediction at horizon $h$. We get:

$$\begin{aligned} E_T(Y_{i,T+h}) &= m + E_T(u_{i,T+h}) \\ &= m + \phi^h E_T[u_{i,T}] \\ &= m + \phi^h Cov[u_{i,T}, (Y_{i,T}, ..., Y_{i,1})'] \left[ V \begin{pmatrix} Y_{i,T} \\ \vdots \\ Y_{i,1} \end{pmatrix} \right]^{-1} \begin{bmatrix} Y_{i,T} - m \\ \vdots \\ Y_{i,1} - m \end{bmatrix}. \end{aligned}$$

Let us denote by $J(\phi)$ the $T \times T$ matrix defined by:

$$J(\phi) = \begin{bmatrix} 1 & \phi & \dots & \phi^{T-1} \\ \phi & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \phi \\ \phi^{T-1} & \dots & \phi & 1 \end{bmatrix},$$

we get:

$$\pi_i(T,h) = m + \phi^h \eta^2 [1, \phi, ..., \phi^{T-1}][\sigma^2 Id + \eta^2 J(\phi)]^{-1} \begin{bmatrix} Y_{i,T} - m \\ \vdots \\ Y_{i,1} - m \end{bmatrix}. \tag{2.7}$$

We note that the pure premium in general depends on the prediction horizon. For a long horizon $[h \to \infty]$ and a regression coefficient $|\phi| < 1$, the pure premium coincides with the population parameter $m$ and the individual contract history provides no useful information. The informational advantage of the insured person depends on the prediction horizon.

When the individual effort $u_{i,T+1}$ is unknown, then even for a short prediction horizon ($h = 1$) and a long contract history ($T \to \infty$) the pure premium $\pi_i(T, 1)$ would not be equivalent to $m + u_{i,T+1}$. For instance, when $\phi = 0$, $u_{i,t}$ and $v_{i,t}$ cannot be distinguished by the econometrician and the observations are not informative to predict $Y_{i,T+1}$. Therefore the informational advantage of the individual also depends on his/her effort modification.

In a neighbourhood of fixed effort, i.e. for $\phi$ close to 1, we can derive the first order expansion of the pure premium [see Appendix 1].

# 3 Nonlinear Panel Model with Heterogeneity

We consider now the extension of linear credibility theory to a nonlinear framework. In the first subsection we describe the case of independent heterogeneity, illustrated by the Poisson-Gamma model with heterogeneity, used in the insurance industry. In the second subsection, we allow for temporal dependence of the heterogeneity components.

### 3.1 Static individual heterogeneity

As in the previous section, the risk variables are denoted by $Y_{i,t}$, whereas the individual heterogeneity is denoted by $u_i$. These variables may be multidimensional. The model is defined under the following set of assumptions:

i) The risk variables are independent conditional on the individual heterogeneity components $u_i,\ i = 1, ..., n$;

ii) The conditional distribution of $Y_{i,t}$ given $u_i,\ i = 1, ..., n$ depends on the conditioning variables through $u_i$ only. Its p.d.f. belongs to a given parametric class of distributions: $f(y|\, u_i; \theta)$ (say), where $\theta$ is a p-dimensional parameter.

iii) The heterogeneity components $u_i,\ i = 1, ..., n$ are independent, identically distributed.

iv) The distribution of $u_i$ is gaussian with zero mean and an identity variance covariance matrix.

Under assumptions i), ii), iii), the heterogeneity components are not uniquely defined. One can easily see that assumption iv) is also satisfied by an appropriately chosen transform of the initial heterogeneity component. Let us assume that the distribution of $u_i$ belongs to a parametric family with the continuous p.d.f. $g(u; \alpha)$, (say). It is always possible to transform $u_i$ into a standard gaussian variable: $u_i = G(u_i^*;\ \alpha)$, (say) where $u_i^* \sim N(0, Id)$ and $G$ is a one to one function. Therefore assumption (iv) is satisfied for $u_i^*$ as new heterogeneity component.

Under assumptions i)-iv) the sequences of risk measures corresponding to two different individuals are independent, and identically distributed. The p.d.f. of $Y_{i,1}, ..., Y_{i,T}$ is given by:

$$l(y_{i,1}, ..., y_{i,T}; \theta) = \int ... \int \prod_{t=1}^{T} f(y_{i,t}|\, u; \theta)\psi(u)du, \tag{3.1}$$

where $\psi$ denotes the p.d.f. of the (multivariate) standard gaussian distribution. The determination of the p.d.f. $l$ may require integration by Monte Carlo methods, except for some specific models, like for example, the Poisson-Gamma model [3].

*Example 3.1 : The Poisson-Gamma model with static heterogeneity.*

In car insurance the risk at period $t$ can be measured by a bidimensional variable, whose first component $Z_{i,t}$ is the number of claims and second component $X_{i,t}$ is the total cost of the claims. A common assumption is a Poisson distribution of $Z$ [ see, Hausman, Hall, Griliches (1984), Gourieroux, Monfort, Trognon(1984), Dionne, Vanasse (1989), Lemaire (1995)], and a gamma-type distribution of $X$ [Gourieroux (1998a)], and usually a heterogeneity factor is introduced for each component of the risk. The distribution of the heterogeneity is usually selected from the conjugate

[3] The case of qualitative data has been discussed, for instance, by Albert, Chib (1993)

families of Poisson and gamma distributions. We denote by $u^*_{1,i}, u^*_{2,i}$ the initial heterogeneity components and by $u_{1,i}, u_{2,i}$ their gaussian transforms. The distributional assumptions of the complete model are given below:

i) The risk variables $(Z_{i,t}, X_{i,t})$, $i, t$ varying, are independent conditional on the heterogeneity components $u^*_{1,i}, u^*_{2,i}$, $i$ varying;

ii) the heterogeneity components $u^*_{1,i}, u^*_{2,i}$, $i$ varying are independent;

iii) the conditional distribution of the count variable $Z_{i,t}$ given the heterogeneity components is the Poisson distribution $\mathcal{P}[u^*_{1,i}\theta_1]$;

iv) the conditional distribution of the cost variable $X_{i,t}$ given $Z_{i,t}$ and the heterogeneity component is the gamma distribution $\gamma[\theta_3 Z_{i,t},\, \theta_2 u^*_{2,i}]$;

v) the marginal distributions of the heterogeneity components are:

$$\begin{aligned} u^*_{1,i} &\sim \gamma[\theta_4, \theta_4], \\ u^*_{2,i} &\sim \gamma[\theta_5, \theta_5]. \end{aligned}$$

Let us denote by $H(.;\theta)$ the quantile function of the $\gamma(\theta, \theta)$ distribution. Their gaussian transformed heterogeneity components are:

$$\begin{aligned} u_{1,i} &= \Psi^{-1}[H(u^*_{1,i};\, \theta_4)], \\ u_{2,i} &= \Psi^{-1}[H(u^*_{2,i};\, \theta_5)], \end{aligned} \tag{3.2}$$

where $\Psi$ is the c.d.f. of the standard normal distribution. Under these specific distributional assumptions we obtain explicit formulas of the conditional distribution of future risks $Y_{i,T+h} = (Z_{i,T+h}, X_{i,T+h})'$ given the contract history $Y_{i,t}, t = 1, ..., T$. In particular the pure premium is independent of the horizon $h$ and is given by: [Gourieroux (1998b)]

$$\begin{aligned} \pi_{i,T}(h) &= E[X_{i,T+h} |\, Z_{i,t}, X_{i,t},\, i = 1, ..., n,\, t = 1, ..., T] \\ &= \frac{\theta_3\theta_1}{\theta_2} \frac{\overline{Z}_{i,T} + \theta_4/T}{\theta_1 + \theta_4/T} \frac{\theta_2 \overline{X}_{i,T} + \theta_5/T}{\theta_3 \overline{Z}_{i,T} + (\theta_5 - 1)/T}. \end{aligned} \tag{3.3}$$

Like in the static linear credibility theory, for a long history ($T \to \infty$) the pure premium no longer depends on the population parameters $\theta$, but instead on the individual historical summary variables $\overline{Z}_{i,T}, \overline{X}_{i,T}$ only.

### 3.2 Dynamic individual heterogeneity

The nonlinear panel model with static heterogeneity is extended by allowing the individual components to be time dependent. To obtain a tractable model with simple interpretation, we consider a gaussian autoregressive heterogeneity of order one. The set of assumptions becomes:

i) The risk variables are independent conditional on the individual heterogeneity components $u_{i,t}$, $i$, $t$ varying;

ii) The conditional distribution of $Y_{i,t}$ given the heterogeneity components depends on the conditioning variables through $u_{i,t}$ only, and belongs to a given parametric class of distributions: $F(y| u_{i,t}; \theta)$, (say);

iii) The individual processes of heterogeneity $(u_{i,t},\ t \geq 0)$, $i$ varying are independent, identically distributed;

iv) They satisfy a gaussian autoregressive model:

$$u_{i,t} = \Phi u_{i,t-1} + (Id - \Phi\Phi')^{\frac{1}{2}} \epsilon_{i,t}, \quad t \geq 1,$$

where $u_{i,0}, \epsilon_{i,t}$, $t \geq 1$ are independent, gaussian variables $N(0, Id)$.

# 4 Statistical Inference

Let us assume that the available information consists of observations $Y_{i,t},\ t = 1, ..., n,\ t = 1, ..., T$. The statistical analysis of a nonlinear panel model with dynamic heterogeneity concerns: i) the estimation of the unknown parameter including the static parameter $\theta$ and the dynamic parameter $\Phi$; ii) the approximation of the unobserved heterogeneity variables $u_{i,t},\ i = 1, ..., n\ t = 1, ..., T$, which may provide a criterion for classification of individuals according to the dynamics of their heterogeneity; iii) the prediction of future values of the process $Y_{i,T+h}$ used for the future risk pricing.

These various problems are rather complex due to the unobservability of the $u$ variables. The unobservability implies that the likelihood function, which underlies the maximum likelihood method, and the conditional expectations, which are used for predicting the $u$ [smoothing] and the future $Y$ [filtering], involves multiple integrals. For instance, the likelihood function conditional on the initial values of heterogeneity is:

$$\prod_{i=1}^{n} \int ... \int \prod_{t=1}^{T} \left\{ f(y_{i,t}| u_{i,t}; \theta) \frac{\exp -\frac{1}{2}(u_{i,t} - \Phi u_{i,t-1})'(Id - \Phi\Phi')^{-1}(u_{i,t} - \Phi u_{i,t-1})}{(2\pi)^{\frac{p}{2}} det(Id - \Phi\Phi')^{\frac{1}{2}}} du_{i,t} \right\},$$

where $p$ is the dimension of the heterogeneity. It is an integral of dimension $pT$ proportional to the number of dated observations. The complexity of this formula suggests to replace analytical computations by simulation based approaches [Gourieroux, Monfort (1996), Richard (1996)]. Fortunately, the nonlinear factor models are easy to simulate.

Let us denote by $\epsilon^s_{i,t},\ i=1,...,n,\ t=1,...,T$ independent drawings from the standard normal distribution. We define recursively simulated heterogeneity values corresponding to the parameter value $\Phi$ by : $u^s_{i,t}(\Phi)=\Phi u^s_{i,t-1}(\Phi)+(Id-\Phi\Phi')^{\frac{1}{2}}\epsilon^s_{i,t},\ i=1,...,n,\ t=1,...,T$. The simulated values of the process corresponding to parameter values $\Phi,\theta$ are found by independent drawings $y^s_{i,t}(\Phi,\theta)$ from the distribution $F[.\,|\,u^s_{i,t}(\Phi);\theta]$. Note that the length of the simulated path can be equal to the length of the observed path, or even longer.

## 4.1 Estimation method

Various simulation based methods can be applied to estimate the parameters $\theta$ and $\Phi$. We describe below two approaches based on the maximum likelihood function and the bounded memory likelihood function.

i) Simulated maximum likelihood function

The idea is to replace the intractable likelihood function by an approximation obtained from simulations. The estimator is defined by:

$$(\hat{\theta}^s_{T,n},\hat{\Phi}^s_{T,n})=\underset{\theta,\Phi}{Argmax}\sum_{i=1}^{n}\log\left[\frac{1}{S}\sum_{s=1}^{S}\prod_{t=1}^{T}f(y_{i,t}|\,u^s_{i,t}(\Phi),\theta)\right], \tag{4.1}$$

where the multidimensional integral in (4.1) is replaced by an empirical average of simulations. The estimator is consistent, when $n$ or $T$ tends to infinity, jointly with the number of replications $S$ [see, Gourieroux, Monfort (1996)].

ii) Bounded memory likelihood method

The simulated maximum likelihood can be quite cumbersome to implement numerically. This is due to the fact that the likelihood function is a nonlinear function of parameters, and that the initial distribution has been replaced by a mixture of conditional distributions. However it is possible to consider a simplified approach based on the marginal likelihood:

$$\prod_{i=1}^{n}\prod_{t=1}^{T}f(y_{i,t};\theta), \tag{4.2}$$

where $f(y;\theta)$ denotes the marginal distribution of $y_{i,t}$, and on the likelihood with memory limited to lag one: [see, Azzalini (1983)]

$$\prod_{i=1}^{n}\prod_{t=1}^{T}f(y_{i,t},y_{i,t-1};\theta,\Phi), \tag{4.3}$$

where $f(y_{i,t},y_{i,t-1};\theta,\Phi)$ denotes the marginal distribution of $y_{i,t},y_{i,t-1}$.

We propose the following two-step approach:

a) First step: estimation of $\theta$ by a marginal maximum likelihood approach.

It is interesting to note that $f(y_{i,t};\theta)$ corresponds to the so-called static model, where $\Phi = 0$. In particular the analytical form of the marginal distribution is often known, like in the example of car insurance. A first step estimator of $\theta$ is:

$$\hat{\theta}_{T,n} = \underset{\theta}{Argmax} \sum_{i=1}^{n} \sum_{t=1}^{T} \log f(y_{i,t};\theta). \tag{4.4}$$

b) Second step: estimation of $\Phi$ by a bounded memory maximum likelihood approach

The joint p.d.f. $f(y_{i,t}, y_{i,t-1};\theta,\Phi)$ is a multidimensional integral, since we have to integrate out the unobservable heterogeneity $u_{i,t}, u_{i,t-1}$. However the dimension of this integral does not depend any more on the number of observations, and is quite low. In the car insurance example it is equal to two, since obviously two integrations can be performed analytically. The second step estimators of $\Phi$ is defined by:

$$\hat{\Phi}_{T,n} = \underset{\Phi}{Argmax} \sum_{i=1}^{n} \sum_{t=1}^{T} \log f(y_{i,t}, y_{i,t-1}, \hat{\theta}_{T,n}, \Phi). \tag{4.5}$$

The asymptotic properties of these estimators can be found in [Azzalini (1983), Gourieroux, Jasiak (1998a)].

## 4.2 Prediction of future risks

As shown in the example of car insurance, the premium is obtained from the conditional distribution of $y_{i,T+1}$ given $\underline{y_{i,T}}$. Moreover the computations described in subsections 2.2 and 3.1 on credibility theory can be performed separately for each individual, and the predictive distribution can be found after having determined the conditional distribution of future heterogeneity $u_{i,T+1}$ given the individual history $\underline{y_{i,T}}$.

We explain below how to obtain these distributions by using approximated bayesian updating equations. Since the computations are performed for a fixed individual, we omit the index $i$.

The bayesian updating equations recursively define the conditional distribution in the following steps:

i) input: the filtering distribution: $l_0(u_t|\underline{y_t})$;

ii) factor predictive distribution: $l_1(u_{t+1}|\underline{y_t}) = \int g(u_{t+1}|u_t) l_0(u_t|\underline{y_t}) du_t$, where $g$ denotes the conditional normal p.d.f. of $u_{t+1}$ given $u_t$;

iii) output: the updated filtering distribution:

$$l_0(u_{t+1}|\underline{y_{t+1}}) = l_1(u_{t+1}|\underline{y_t}) f(y_{t+1}|u_t) \left[\int l_1(u_{t+1}|\underline{y_t}) f(y_{t+1}|u_{t+1}) du_{t+1}\right]^{-1}.$$

We now replace in this recursive scheme all distributions by discrete approximations derived by simulations. The procedure is outlined below:

i) input: an approximated filtering distribution: $\hat{l}_0(u_t|\underline{y_t})$, where $\hat{l}_0$ is the empirical distribution corresponding to some values $u_t^{0,1}, ..., u_t^{0,S}$;

ii) approximated factor predictive distribution: $\hat{l}_1(u_{t+1}|\underline{y_t})$.

To each value $u_t^{0,s}$, we associate a value $u_t^{1,s}$ drawn from the distribution $g(.|u_t^{0,s})$. The approximation $\hat{l}_1(.|\underline{y_t})$ is the empirical distribution of $u_{t+1}^{1,1}, ..., u_{t+1}^{1,S}$.

iii) output: an approximated updated filtering distribution $\hat{l}_0(u_{t+1}|\underline{y_{t+1}})$.

Since the filtering distribution $l_0(u_{t+1}|\underline{y_{t+1}})$ is approximately proportional to $\hat{l}_1(u_{t+1}|\underline{y_t})f(y_{t+1}|u_t)$ whose analytical form is known, we apply a standard acceptance-rejection method based on the latter function [see, e.g. Richard (1996)].

Let us compute $f_{t+1}^{\max} = \max_{s=1,...,S} f(y_{t+1}|u_{t+1}^{1,s})$. Then, for each $s$ we consider a first draw from $u_{t+1}^{1,1}, ..., u_{t+1}^{1,S}$ yielding $u_{1,s}^*$, say. This draw is retained with probability $\frac{f(y_{t+1}|u_{1,s}^*)}{f_{t+1}^{\max}}$, rejected otherwise. If it is rejected, we consider a second draw in $\hat{l}_1$, getting $u_{2,s}^*$, say, and reapply a selection with probability $\frac{f(y_{t+1}|u_{2,s}^*)}{f_{t+1}^{\max}}$, and so on, until we obtain a first accepted value $u_{t+1}^{1,s}$. The approximated distribution $\hat{l}_0(.|\underline{y_{t+1}})$ is the empirical distribution of $u_{t+1}^{0,1}, ..., u_{t+1}^{0,S}$.

The algorithm can be continued to provide an approximated predictive distribution:

iv) approximated predictive distribution of $y_{t+1}$ : $\hat{l}_1^P(y_{T+1}|\underline{y_T})$.

To each value $u_{T+1}^{1,s}$ corresponds a drawing $y_{T+1}^s$ in the distribution $f(y_{T+1}|u_{T+1}^{1,s})$. The predictive distribution is approximated by the empirical distribution of $y_{T+1}^s$, $s = 1, ..., S$.

### 4.3 Reconstruction of past heterogeneity

It is also interesting to develop a simulation based approach to determine the conditional distribution of past heterogeneity values $u_1, ..., u_T$ given the observations $y_1, ..., y_T$. Drawings in the conditional distribution $l(u_1, ..., u_T|\underline{y_T})$, say can be obtained from a Gibbs sampler [see Chib (1996) for applications to panel data models]. This procedure consists of the following steps:

i) We select initial values $u_1^0, ..., u_T^0$ for the past heterogeneity.

ii) Let us denote by $u_1^{j-1}, ..., u_T^{j-1}$ the values available at the beginning of step $j$. Then we draw recursively the next values $u_t^j$, $t$ varying, in the conditional distribution of $u_t$ given $u_1 = u_1^j, ..., u_{t-1} = u_{t-1}^j, u_{t+1} = u_{t+1}^{j-1}, ..., u_T = u_T^{j-1}, \underline{y_T}$. When $j$ is sufficiently large $u_1^j, ..., u_T^j$ is close to a drawing from the conditional distribution $l(u_1, ..., u_T|\underline{y_T})$.

The advantage of Gibbs sampling is that only drawings from a conditional distribution of the type $l(u_t|u_1, ..., u_{t-1}, u_{t+1}, ..., u_T, \underline{y_T})$ have to be performed. Such drawings are easy to perform since the conditional distribution is proportional to:

$$l(u_t|u_1, ..., u_{t-1}, u_{t+1}, ...u_T, \underline{y_T}) \quad = \quad l(u_t|u_{t-1}, u_{t+1}, y_t)$$

$$\propto \quad g(u_t|u_{t-1})g(u_{t+1}|u_t)f(y_t|u_t),$$

which has a simple analytical form. It allows us to perform drawings using a standard acceptance-rejection method as shown in the previous subsection.

# 5 Impulse response functions

## 5.1 Definition

The impact of heterogeneity on the observed variables can be analyzed through appropriate impulse response functions. Contrary to the suggestion by Gallant, Rossi, Tauchen (1993) to consider only consequences of a shock to $y_T$, say, on the future values $y_{T+1}, ..., y_{T+H}$, it seems more appropriate in our framework to distinguish various basic shocks, which may influence the $y$ variable. They include the components of heterogeneity factors and the residual randomness associated with the conditional distribution $f(y_t|u_t)$, encountered in the static case $\Phi = 0$. We are essentially concerned in this section about the transmission of transitory or permanent shocks to the innovation term of the heterogeneity factor dynamics at the individual level [4].

In this experiment we fix at a constant level the individual past heterogeneity factors and the observables until date $T$, or alternatively the most recent realizations $u_T = u_T^0, y_T = y_T^0$, say. The future values of the observable process depend both on the sequence of future heterogeneity shocks $\epsilon_{T+1}, ..., \epsilon_{T+H}$, and the sequence of conditional shocks $\eta_{T+1}, ..., \eta_{T+H}$, corresponding to the drawings of $y_{T+h}$ given $u_{T+h}$. These conditional shocks can be represented by independent variables with identical standard gaussian distributions [5]. The future value $y_{T+H}$ can be written as a deterministic function of these shocks and the initial state values:

$$y_{T+H} = a_H(u_T^0, y_T^0; \epsilon_{T+1}, \eta_{T+1}, ..., \epsilon_{T+H}, \eta_{T+H}). \tag{5.1}$$

Let us now assume the occurrence of some deterministic shocks to the innovation term of the heterogeneity process.

If this shock is of magnitude $\delta$ and transitory, occuring at date $T+1$, the future values become:

$$y_{T+H}^{tr}(\delta) = a_H(u_T^0, y_T^0; \epsilon_{T+1} + \delta, \eta_{T+1}, ..., \epsilon_{T+H}, \eta_{T+H}). \tag{5.2}$$

If this shock is of magnitude $\delta$ and permanent, the future values become:

$$y_{T+H}^{p}(\delta) = a_H(u_T^0, y_T^0; \epsilon_{T+1} + \delta, \eta_{T+1}, \epsilon_{T+2} + \delta, \eta_{T+2}, ..., \epsilon_{T+H} + \delta, \eta_{T+H}). \tag{5.3}$$

[4] As before, the individual index is omitted

[5] this is always possible using the standard inversion technique

To analyse the effects of shocks we can consider the joint distribution of the benchmark variable $y_{t+H}$ and the perturbed variable $y^{tr}_{T+H}(\delta)$ or $y^{p}_{T+H}(\delta)$ conditional on $u^0_T, y^0_T$, and study how it varies with the magnitude of shock $\delta$ and the horizon $H$ [Gourieroux, Jasiak (1998)].

In practice, these joint distributions are approximated by simulations, i.e. by the empirical distributions of: $(y^s_{T+H}, y^{tr,s}_{T+H}(\delta))$ or $(y^s_{T+H}, y^{p,s}_{T+H}(\delta))$, where the simulated $y$ values are obtained by replacing the terms $\epsilon_{T+1}, ..., \epsilon_{T+H},\ \eta_{T+1}, ..., \eta_{T+H}$ in equations (5.1), (5.2), (5.3) by independent drawings $\epsilon^s_{T+1}, ..., \epsilon^s_{T+H},\ \eta^s_{T+1}, ..., \eta^s_{T+H},\ s = 1, ..., S$ in the standard gaussian distribution.

It may also be of interest to describe these distributions and compare some selected moments $E\, b(y_{T+H})$ and $E\, b(y^{tr}_{T+1}(\delta))$, where $b$ is given, or their simulated counterparts [see, e.g. Koop, Pesaran, Potter (1996)].

## 5.2 Impulse response analysis of the Poisson-Gamma model with dynamic heterogeneity

i) The model

We now consider the impulse response analysis for the Poisson-Gamma model with dynamic heterogeneity. The observable variable is bidimensional with two heterogeneity factors. We have: $Y_t = (X_t, Z_t)$, where $X_t$ is the annual cost of accidents and $Z_t$ the annual number of accidents. The conditional distributions are:

$Z_t | u^*_t, v^*_t \sim \mathcal{P}[u^*_t \theta_1]$,

$X_t | Z_t, u^*_t, v^*_t \sim \gamma(\theta_3 Z_t,\ \theta_2 v^*_t)$,

$u^*_T = H^{-1}(\Psi(u_t), \theta_4)$,

$v^*_t = H^{-1}(\Psi(v_t), \theta_5)$,

$$\begin{pmatrix} u_t \\ v_t \end{pmatrix} = \Phi \begin{pmatrix} u_{t-1} \\ v_{t-1} \end{pmatrix} + (Id - \Phi\Phi')^{1/2} \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}.$$

The parameters $\theta_1$ and $\theta_4$ are set at $\theta_1 = 0.2$ and $\theta_4 = 0.1$. This implies that the expected number of accidents is $EZ_t = \theta_1 = 20\%$, while its variance $VZ_t = V(\theta_1 u^*_t) + E(\theta_1 u^*_t) = \theta_1[1 + \frac{\theta_1}{\theta_4}]$ allows for an overdispersion of $\theta_1/\theta_4 = 2$.

The parameters $\theta_3$ and $\theta_2$ are set at $\theta_3 = 100, \theta_2 = 0.1$. In absence of heterogeneity $v^*_t = 1 (\Leftrightarrow \theta_5 = +\infty)$, the expected cost per accident is $\theta_3/\theta_2 = 1000\$$, whereas the corresponding variance is $\theta_3/\theta_2^2 = 10000$. The magnitude of the heterogeneity $v^*_t$ is fixed at the same level as for $u^*_t : \theta_5 = 0.1$.

Finally we introduce the independent dynamics of heterogeneity factors:

$$\begin{aligned} u_t &= 0.8 u_{t-1} + \sqrt{1 - 0.8^2}\epsilon_{1t}, \\ v_t &= 0.5 v_{t-1} + \sqrt{1 - 0.5^2}\epsilon_{2t}. \end{aligned}$$

ii) The experiment

The initial values of the heterogeneity factors are fixed at their marginal expected values: $u_T = v_T = 0$. In the next step, we perform the following experiments to examine consequences of:

experiment I: a transitory shock of $\delta = 1$ on $\epsilon_{1t}$.

experiment II: a transitory shock of $\delta = -1$ to $\epsilon_{1t}$

experiment III: a transitory shock of $\delta = 1$ to $\epsilon_{2t}$.

experiment IV: a transitory shock of $\delta = -1$ to $\epsilon_{2t}$.

experiment V: a permanent shock of $\delta = 1$ to $\epsilon_{1t}$.

experiment VI: a permanent shock of $\delta = 1$ to $\epsilon_{2t}$.

We focus on the prediction of the total cost at horizon $H$. Figure 5.1 displays scatters representing joint distributions of the cost and perturbed cost for the first experiment at horizons $H = 1, 5, 10$. The perturbation results from a transitory shock to the occurrence heterogeneity,

insert Figure 5.1: Scatterplots of costs and perturbed costs;

We note that under a negative [resp. positive] shock the observations tend to cluster above [resp. below] the 45 degree line. They approach the 45 degree line when the horizon increases. Since the dynamics is nonlinear, we do not observe symmetric effects of positive and negative shocks of the same magnitude. In particular, responses to positive shocks are more persistent than responses to negative shocks.

In the insurance analysis we are mainly interested in finding the expected total cost and its variance. We provide in Figures 5.2 the expected costs as functions of the horizon for the experiments outlined above, and in Figures 5.3 the associated variances.

In each Figure, we plot the baseline corresponding to the absence of shock, and the moments perturbed by negative and positive shocks, respectively. We find that shocks to both heterogeneity factors impact the first and second order moments of the total cost, with effects of shocks hitting the occurrence heterogeneity being stronger. The asymmetric responses to positive and negative shocks are also observed and even more pronounced in the experiment involving the permanent shocks.

insert Figure 5.2: Comparison of expected costs

insert Figure 5.3: Comparison of the cost variabilities.

# Appendix

Linear credibility theory in a neighbourhood of static individual heterogeneity.

We perform the first order expansion of the pure premium given in (2.7), when $\phi \approx 1$. We denote by:

$$J(\phi) = J - (1-\phi)K + o(1-\phi),$$

the first order expansion of the matrix $J(\phi)$. The $J$ matrix is equal to $J = ee'$, where $e$ is the $T$-dimensional vector with unitary elements. The $K$ matrix is:

$$K = \begin{bmatrix} 0 & 1 & 2 & \dots & T-1 \\ 1 & 0 & \ddots & \ddots & \vdots \\ 2 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ T-1 & \dots & \dots & 1 & 0 \end{bmatrix},$$

and its first column will be denoted by $k$. Then we get:

$$\begin{aligned}
\pi_i(T,h) &= m + [1-h(1-\phi)]\eta^2[e' - (1-\phi)k'] \\
&\quad [\sigma^2 Id + \eta^2 J - (1-\phi)\eta^2 K]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + o(1-\phi) \\
&= m + [1-h(1-\phi)]\eta^2[e' - (1-\phi)k'] \\
&\quad \{Id - (1-\phi)\eta^2(\sigma^2 Id + \eta^2 J)^{-1}K\}^{-1}[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + o(1-\phi) \\
&= m + [1-h(1-\phi)]\eta^2[e' - (1-\phi)k'] \\
&\quad \{Id + (1-\phi)\eta^2(\sigma^2 Id + \eta^2 J)^{-1}K\}[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + o(1-\phi) \\
&= m + \eta^2 e'[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} \\
&\quad -(1-\phi)h\eta^2 e'[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} - (1-\phi)\eta^2 k'[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} \\
&\quad +(1-\phi)\eta^4 e'[\sigma^2 Id + \eta^2 J]^{-1}K[\sigma^2 Id + \eta^2 J]^{-1} \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + o(1-\phi).
\end{aligned}$$

Moreover we have:

$$[\sigma^2 Id + \eta^2 J]^{-1} = \frac{1}{\sigma^2}[Id - \frac{\eta^2}{\sigma^2+\eta^2 T}J],$$
$$[\sigma^2 Id + \eta^2 J]^{-1}e = \frac{1}{\sigma^2+\eta^2 T}e.$$

We deduce that

$$\begin{aligned}
\pi_i(T,h) &= m + \frac{\eta^2}{\sigma^2+\eta^2 T}e' \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} - (1-\phi)h\frac{\eta^2}{\sigma^2+\eta^2 T}e' \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} \\
&\quad -(1-\phi)\frac{\eta^2}{\sigma^2}k'\left[Id - \frac{\eta^2}{\sigma^2+\eta^2 T}ee'\right] \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} \\
&\quad +\frac{(1-\phi)}{\sigma^2}\frac{\eta^4}{\sigma^2+\eta^2 T}e'K\left[Id - \frac{\eta^2}{\sigma^2+\eta^2 T}ee'\right] \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + o(1-\phi) \\
&= m + \frac{\eta^2 T}{\sigma^2+\eta^2 T}(\overline{Y}_{i,T}-m) - (1-\phi)h\frac{\eta^2 T}{\sigma^2+\eta^2 T}(\overline{Y}_{i,T}-m) \\
&\quad -(1-\phi)\frac{\eta^2}{\sigma^2}k' \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} + (1-\phi)\frac{\eta^4 T}{\sigma^2(\sigma^2+\eta^2 T)}\frac{T(T-1)}{2}(\overline{Y}_{i,T}-m) \\
&\quad +\frac{1-\phi}{\sigma^2}\frac{\eta^2}{\sigma^2+\eta^2 T}e'K \begin{bmatrix} Y_{i,T}-m \\ \vdots \\ Y_{i,1}-m \end{bmatrix} - \frac{(1-\phi)\eta^6 T}{\sigma^2(\sigma^2+\eta^2 T)^2}e'Ke(\overline{Y}_{i,T}-m) + o(1-\phi).
\end{aligned}$$

FIG 5.1 : Scatterplots of costs and perturbed costs

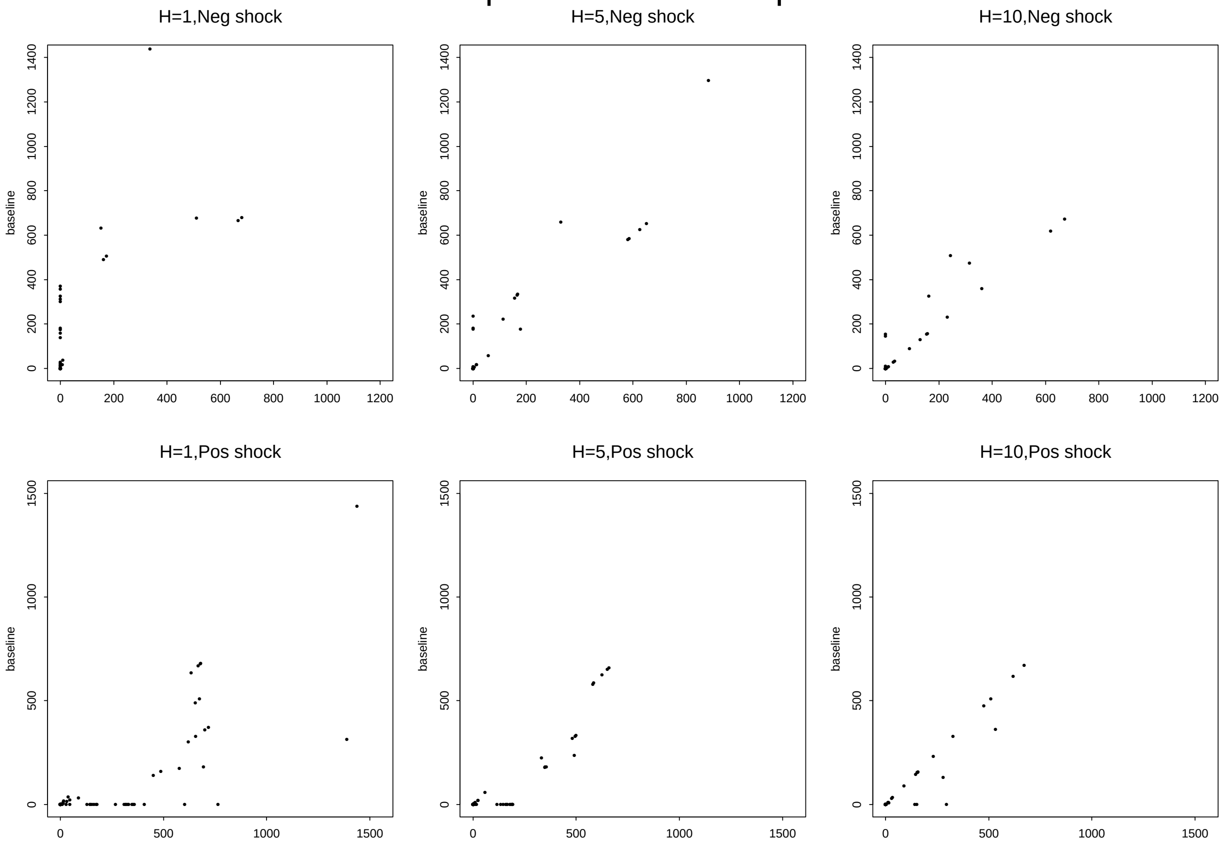


FIG 5.2 : Comparison of expected costs

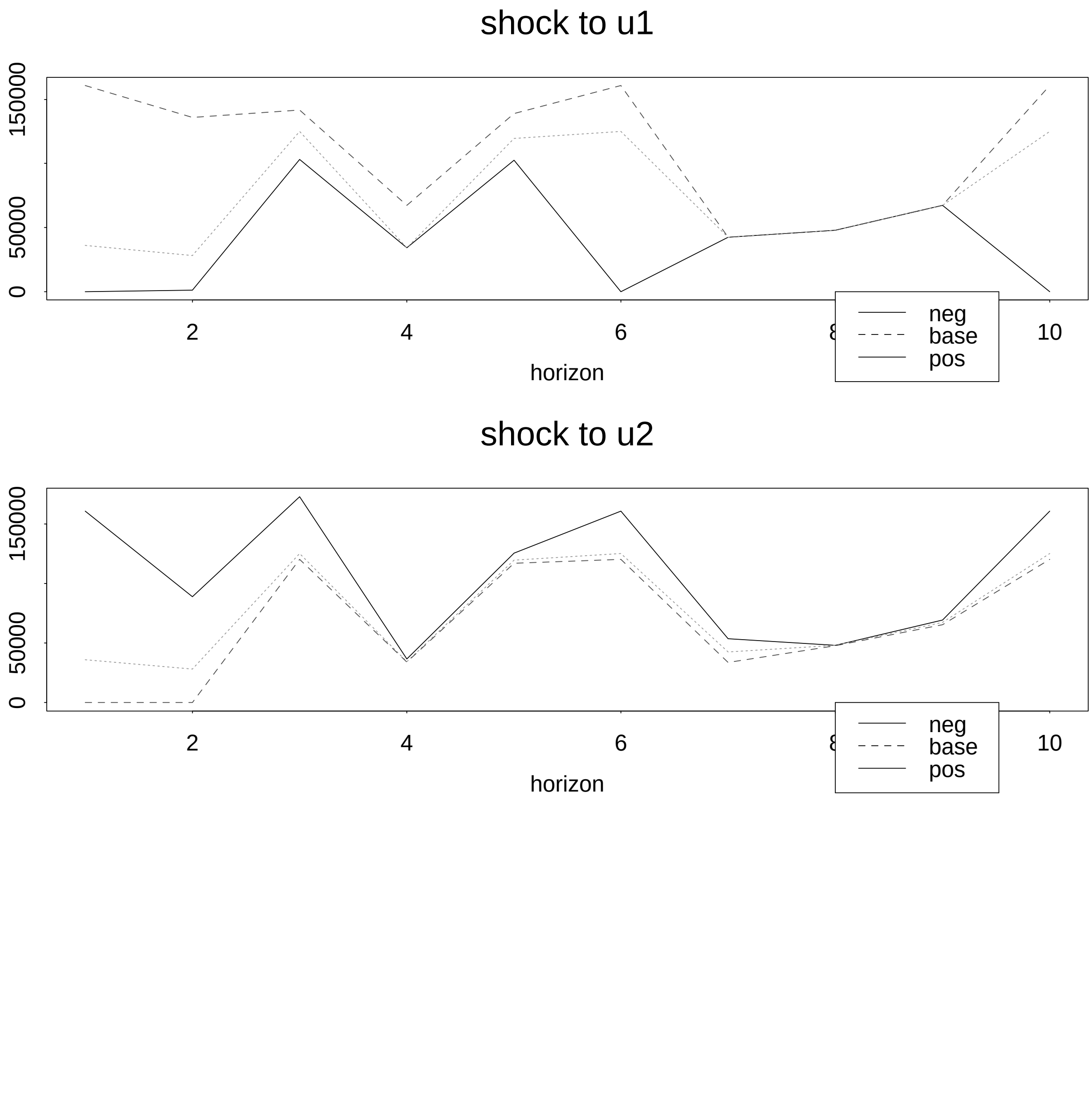


FIG 5.3 : Comparison of st.dev. of costs

shock to u1

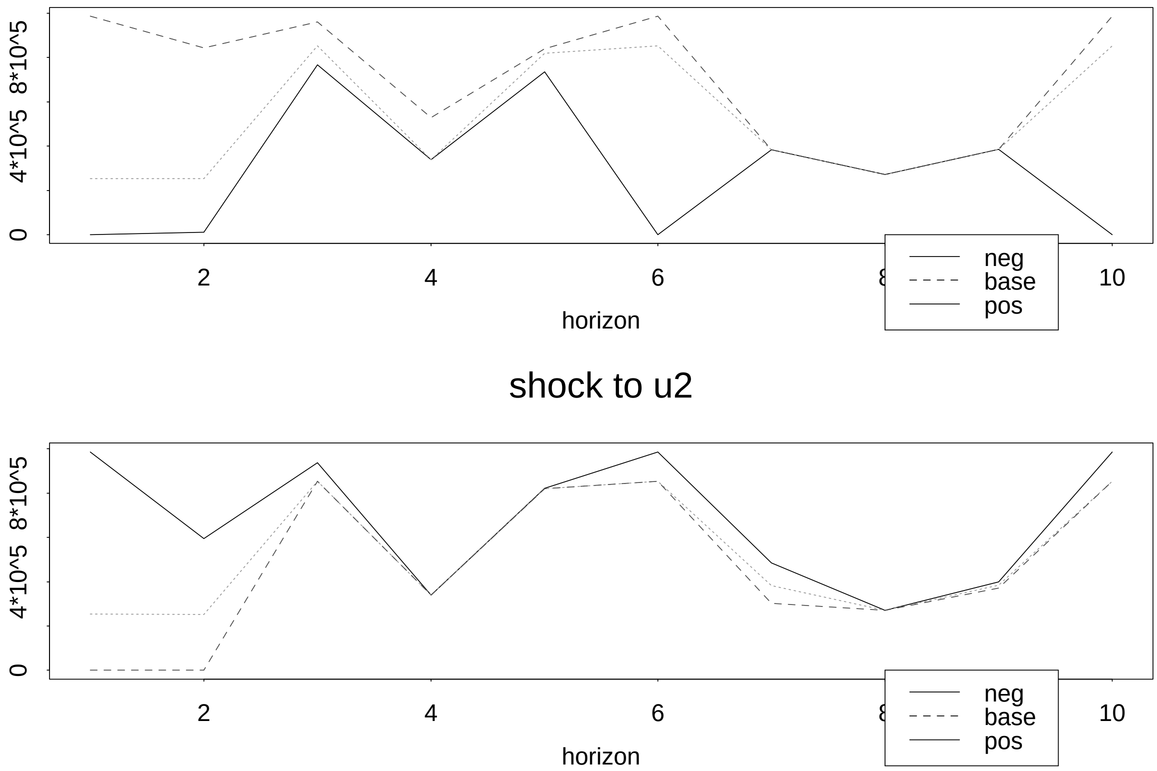
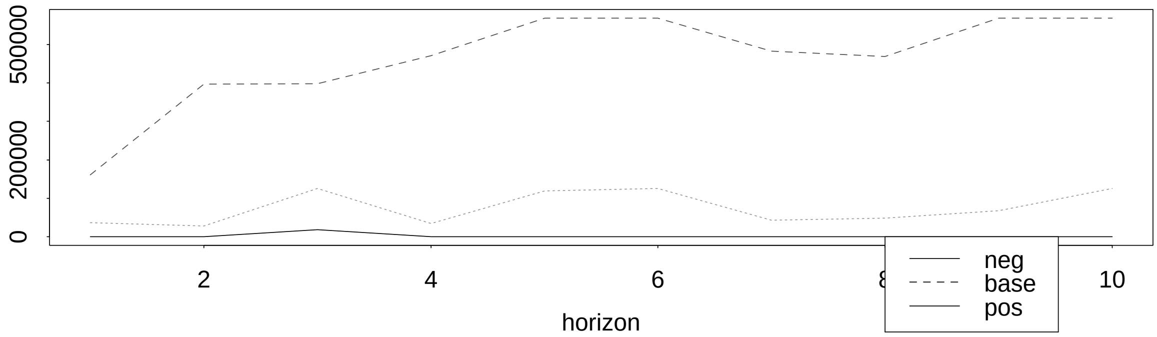


FIG 5.2a : Comparison of expected costs

permanent shock to u1




permanent shock to u2

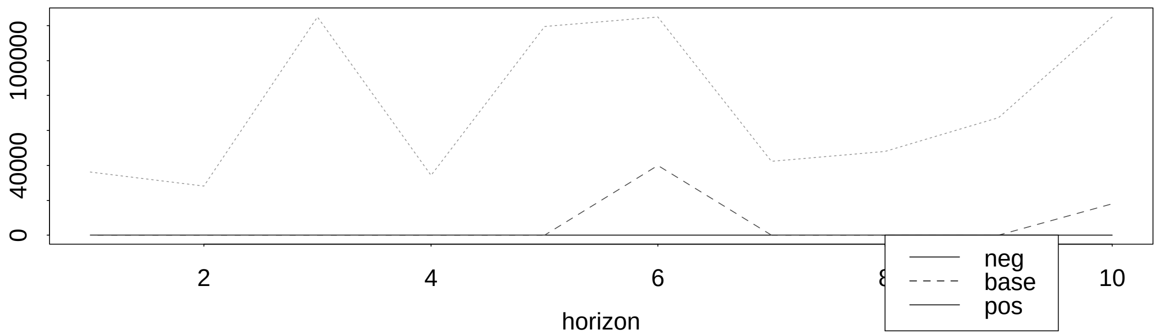
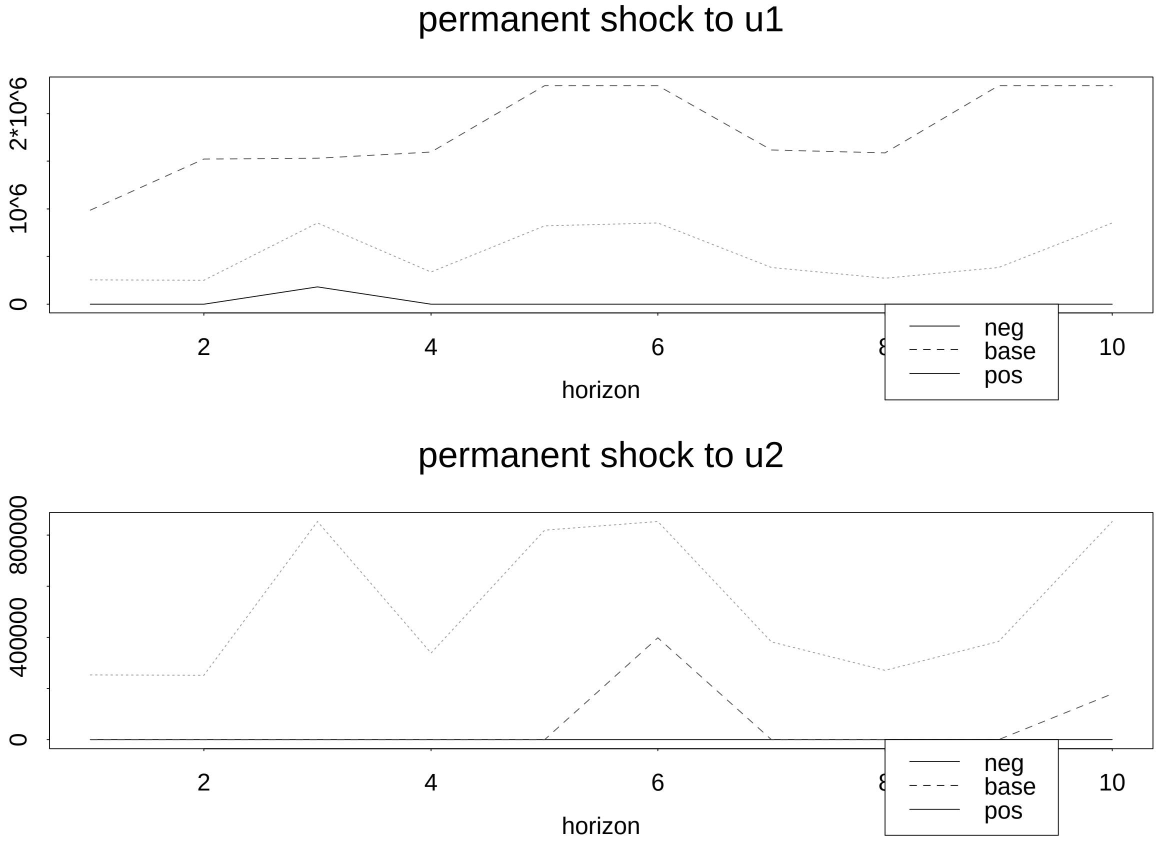


FIG 5.3a : Comparison of std.dev. of costs



## References

[1] Albert, J., and S. Chib (1993): "Bayesian Analysis of Binary and Polytochomous Response Data", Journal of the American Statistical Association, 88, 669-679.

[2] Azzalini, A. (1983): "Maximum Likelihood Estimation of Order m for Stationary Stochastic Processes", Biometrika, 70, 381-387.

[3] Bailey, A. (1945): "A Generalized Theory of Credibility", Proc. Cas. Act. Soc., 32, 13-20.

[4] Bailey, A. (1950): "Credibility Procedures, Laplace's Generalization of Bayes Rule and the Combination of Collateral Knowledge with Observed Data", Proc. Cas. Act. Soc., 37, 7-23.

[5] Balestra, P (1996): "Introduction to Linear Models for Panel Data", in Matyas, L., and P. Sevestre ed., "The Econometrics of Panel Data", 23-33, Kluwer.

[6] Balestra, P., and M. Nerlove (1966): "Pooling Cross-Section and Time Series Data in the Estimation of a Dynamic Model: The Demand for Natural Gas", Econometrica, 34, 585-612.

[7] Balestra, P., and M. Nerlove (1996): "Formulation and Estimation of Econometric Models for Panel Data", in Matyas, L., and P. Sevestre ed., "The Econometrics of Panel Data", 3-22, Kluwer.

[8] Buhlmann, H. (1967): "Experience Rating and Credibility: I", ASTIN Bulletin, 4, 199-207.

[9] Buhlmann, H. (1969): "Experience Rating and Credibility: II", ASTIN Bulletin, 5, 157-165.

[10] Buhlmann, H., and E. Straub (1970): "Glaubwurdigkeit fur Schadensatze", Mitteilungen der Schweizerischen Vereinigung den Versicherungsmathematiken, 111-133.

[11] Chib, S. (1996): "Inference in Panel Data Models via Gibbs Sampling", in Matyas, L., and P. Sevestre ed., "The Econometrics of Panel Data", 639-650, Kluwer.

[12] Dannenburg, D. (1994): "Some Results on the Estimation of the Credibility Factor in the Classical Buhlmann Model", Insurance: Mathematics and Economics, 14, 39-50.

[13] Dionne, G., and C. Vanasse (1989) : "A Generalization of Actuarial Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component", Actuarial Bulletin, 19, 199-212.

[14] Gallant, A., Rossi, P. ,and G. Tauchen (1993): "Nonlinear Dynamic Structures", Econometrica, 61, 871-908.

[15] Goovaerts, M., Bauwelinck, T., and C. Stoop (1989): "The Practical Application of Credibility Theory", Insurance: Mathematics and Economics, 8, 77-96.

[16] Gourieroux, C. (1997): "Econometrics of Insurance", Geneva Papers for Risk and Insurance, forthcoming.

[17] Gourieroux, C. (1998a): "Statistique de l'assurance", 300p, forthcoming, Economica.

[18] Gourieroux, C., and J. Jasiak (1998a): "Dynamic Factor Models", CREST, D.P.

[19] Gourieroux, C., and J. Jasiak (1998b): "Nonlinear ARMA Models", CREST, D.P.

[20] Gourieroux, C., and A. Monfort (1996): "Simulation Based Econometric Methods", Oxford University Press.

[21] Heilmann, W. (1989): "Decision Theoric Foundations of Credibility Theory", Insurance: Mathematics and Economics, 8, 77-96.

[22] Jewell, W. (1983): "A Survey of Credibility Theory", Operation Research Center, 76.31, University of California, Berkeley.

[23] Koop, G., Pesaran, H., and S. Potter (1996): "Impulse Response Analysis in Nonlinear Multivariate Models", Journal of Econometrics, 74, 119-147.

[24] Lemaire, J. (1995): "Bonus-Malus Systems in Automobile Insurance", Kluwer Academic Publishers.

[25] Loubergé, H. (1998): "Risk and Insurance Economics 25 Years After", the Geneva Papers on Risk and Insurance, 23, 540-567.

[26] Matyas, L., and P. Sevestre (1996): "The Econometrics of Panel Data", second edition, Kluwer.

[27] Richard, J.F. (1996): "Simulation Techniques", in Matyas, L., and P. Sevestre,ed., "The Econometrics of Panel Data", 613-638, Kluwer.

[28] Sundt, B. (1989): "Bonus Hunger and Credibility Estimators with Geometric Weights", Insurance: Mathematics and Economics, 8,119-126.