Bonus-Malus Scales Models for Predictive Modeling

CAS Ratemaking, Product and Modeling Seminar

Prof. Jean-Philippe Boucher, ACIA, Ph.D.
Co-operators Chair in Actuarial Risk Analysis
Université du Québec à Montréal (UQAM)

March 15th, 2023

UQÀM | co-operators

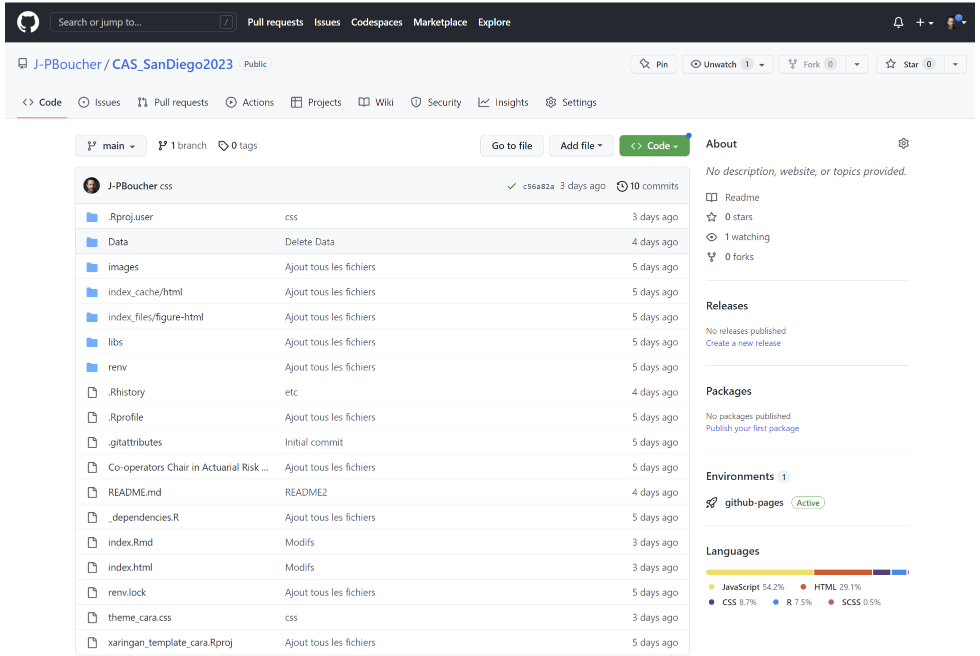
Co-operators Chair in Actuarial Risk Analysis

Scripts and dataframe available on Github

On my github page, you can find:

- this presentation;
- all R script codes used;
- the dataframe *df2.Rda*.

https://github.com/J-PBoucher



R Preamble (to replicate the results)

```
library(tidyverse)
library(xaringan)
library(xaringanthemer)
library(kableExtra)
library(DT)
library(dplyr)
library(ggplot2)
library(kableExtra)
library(scales)
library(MASS)
library(gamlss)

load('Data/df2.Rda')
```

Summary of the presentation

Part I - Ratemaking with Cross-Section Data

- Basic Count Distributions;
- Credibility Models and Predictive Ratemaking;
- Bonus-Malus Scales Models.

Part II - Ratemaking with Panel Data

- Families of Count Distributions;
- Observed Predictive Premiums;
- Bonus-Malus Scales Models Revisited.

Part III - Actual Challenges

- Entry levels and new insureds;
- Penalties and *a priori* risks.

Reference

The first part of the presentation is based on Sections 1-3 of:

J.-P. Boucher (2022). Bonus-Malus Scale Models: Creating Artificial Past Claims History. *Annals of Actuarial Science*, 1-27.

Annals of Actuarial Science (2022), 1–27
doi:10.1017/S1748499522000100

Institute and Faculty of Actuaries

ORIGINAL RESEARCH PAPER

Bonus-Malus Scale models: creating artificial past claims history

Jean-Philippe Boucher

Chaire Co-operators en analyse des risques actuariels, Departement de mathematiques, Université du Québec à Montréal, Montreal, Canada
E-mail: boucher.jean-philippe@uqam.ca

(Received 04 January 2022; revised 09 June 2022; accepted 24 June 2022)

Abstract

In recent papers, Bonus-Malus Scales (BMS) estimated using data have been considered as an alternative to longitudinal data and hierarchical data approaches to model the dependence between different contracts for the same insured. Those papers, however, did not discuss in detail how to construct and understand BMS models, and many of the BMS's basic properties were not discussed. The first objective of this paper is to correct this situation by explaining the logic behind BMS models and by describing those properties. More particularly, we will explain how BMS models are linked with simple count regression models that have covariates associated with the past claims experience. This study could help actuaries to understand how and why they should use BMS models for experience rating. The second objective of this paper is to create artificial past claims history for each insured. This is done by combining recent panel data theory with BMS models. We show that this addition significantly improves the prediction capacity of the BMS and provides a temporary solution for insurers who do not have enough historical data. We apply the BMS model to real data from a major Canadian insurance company. Results are analysed deeply to identify specific aspects of the BMS model.

Data used

Farm insurance database (in the published paper)

- We used farm insurance data from a major insurance company in Canada;
- We were able to use contracts from 2014 to 2019;
- Past claims from 1999 to 2014 were available.

Farm data cannot be shared...

- Instead, we illustrate our models with fictive car insurance data.
- Using the dataframe *df2.Rda*;
- Possible to replicate the results of this presentation;
- Available on my *github* page (reference at the end).

Part I - Ratemaking with Cross-Section Data

- Basic count distributions
- Credibility Models and Predictive Ratemaking
- Bonus-Malus Scales Models

Part II - Ratemaking with Panel Data

- Families of Count Distributions
- Observed Predictive Premiums
- Bonus-Malus Scales Models Revisited

Part III - Actual Challenges

- Entry levels and new insureds
- Penalties and *a priori* risks

Classic insurance database

policy_no	veh.num	renewal_date	start_date	end_date	risk_expo	freq_payment	year_veh	sex	year_b
8352232	1	2018-03-23	2017-03-23	2018-03-22	1	1	2013	M	
8045623	1	2018-07-31	2017-07-31	2018-07-30	1	1	2012	M	
8137137	1	2018-02-12	2017-02-12	2018-02-11	1	12	2012	F	
6590411	1	2012-05-08	2011-05-08	2012-05-07	1	12	2007	M	
6652522	1	2014-05-24	2013-05-24	2014-05-23	1	1	2009	F	
6926801	1	2017-06-01	2016-06-01	2017-05-31	1	1	2009	M	

Summary of the dataset

Basic statistics

Nb. of Claims	Nb. of obs.	% of obs.	Total exposition	% of exposition
0	105,050	83.8%	100,614	83.5%
1	17,734	14.1%	17,246	14.3%
2	2,339	1.9%	2,305	1.9%
3	240	0.2%	237	0.2%
4	25	0.0%	25	0.0%
5	1	0.0%	1	0.0%
6	1	0.0%	1	0.0%

Mean	0.185
Variance	0.203

Available covariates

Many fictional covariates are available in the dataframe. For illustration, we will however only focus on 4 (fictional) covariates to model the number of claims:

Column	Values
car_color	("Other", "Red")
territory	("Rural", "Suburban", "Urban")
language	("English", "French")
food	("Other", "Vegan", "Vegetarian")

Basis of covariates selection

Possible techniques

- Minimum-bias techniques (...old);
- Generalized linear models and GLM-net (Ridge and Lasso);
- Random Forests;
- Neural Networks;
- etc.

Literature review (from actuarial sciences):

- Denuit, M., Hainaut, D. & Trufin, J. (2019), Springer Nature:
    - *Effective statistical learning methods for actuaries I: GLMs and Extensions*,
    - *Effective statistical learning methods for actuaries II: Tree-Based Methods and Extensions*,
    - *Effective statistical learning methods for actuaries III: Neural Networks and Extensions*,
- Wüthrich, M. V., & Merz, M. (2023). *Statistical foundations of actuarial learning and its applications*.

Prior ratemaking

Poisson distribution

Commonly, the starting point for the modeling the number of claims is the Poisson distribution:

$$\Pr[N_i = n_i | \boldsymbol{X}_i] = \frac{\lambda_i^{n_i} e^{-\lambda_i}}{n_i!}, \text{ with } \lambda_i = \exp(\boldsymbol{X}_i' \boldsymbol{\beta}).$$

With $E[N_i] = \lambda_i$, this form of ratemaking is usually called *a priori* ratemaking. In this framework, the actuary does not consider the past claim experience of the insureds.

Alternatives

To correct the equidispersion of the Poisson or other problems, the most popular alternatives to the Poisson are:

- Negative binomial (NB2 or NB1);
- Poisson-inverse gaussian (PIG2 or PIG1);
- Poisson-lognormal (PLN2 or PLN1);
- Zero-Inflated distributions.

Predictive ratemaking

Conditional expected value

The insurer is also interested in a premium that considers past contracts:

$$E[N_{i,T}|n_{i,1}, \ldots, n_{i,T-1}, \boldsymbol{X}_{i,(1:T-1)}].$$

Problem with cross-section data

- We suppose an independance between each line of the dataset;
- We do not directly observe the claim experience of an insured for his 2nd, 3rd, ..., contracts.

Classic assumption (Bühlmann, 1967)

We suppose that each insured has his own random heterogeneity component (usually noted Θ) that affects all his insurance contracts.

Gamma heterogeneity

If we suppose that $N_i|\Theta = \theta \sim Poisson(\lambda_i \theta)$, with $\Theta \sim gamma(\alpha, \tau = \alpha)$, we have:

$$\Pr[N_i = n] = \int_0^{\infty} \frac{(\lambda_i \theta)^n e^{-\lambda_i \theta}}{n!} \frac{\alpha^{\alpha}}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\alpha\theta} d\theta$$

Negative binomial 2 distribution

It can be shown that the Poisson-gamma leads to the NB2, having the following probability function:

$$\Pr[N_i = n] = \binom{\alpha + n - 1}{n} \left(\frac{\alpha}{\lambda_i + \alpha}\right)^{\alpha} \left(\frac{\lambda_i}{\lambda_i + \alpha}\right)^{n}$$

Moments

The NB2 has and expected value of $E[N_i] = \lambda_i$ and a variance $Var[N_i] = \lambda_i + \frac{\lambda_i^2}{\alpha}$, which means that the NB2 allows for overdispersion.

Bayesian approach

The addition of an heterogeneity term leads to the famous credibility models of Buhlmann or Buhlmann-Straub (exam C, or STAM).

Predictive premium

It becomes possible to express the predictive premium based on the past number of claims n, and the past values of λ :

$$E[N_{i,T}|n_{i,1},\ldots,n_{i,T-1},\boldsymbol{X}_{i,T}] = \lambda_{i,T}\frac{\alpha+\sum_{t=1}^{T-1} n_{i,t}}{\alpha+\sum_{t=1}^{T-1} \lambda_{i,t}}$$

Even if the actuary cannot directly observe the average predictive value with the data, he is able to compute predictive premiums at the cost of making the constact random effect assumption.

R scripts: Poisson GLM

Split data

```
db.train <- df2 %>% filter(Type=='TRAIN')
db.test <- df2 %>% filter(Type=='TEST')
```

Poisson GLM

```
score.nbclaim <- as.formula(NbClaims ~ car_color + need_glasses + territory + language + food + offs
Poisson    <- glm(score.nbclaim, family=poisson(link=log), data=db.train)
```

R scripts: Negative Binomial 2

Packages to use

You can directly estimate the parameters by maximum likelihood by maximising the log-probability, or you can use R packages. The *MASS* package (or the *gamlss*, for example) can be used to estimate the parameters of a NB2 distribution:

```
library(MASS)
```

```
nb2.MASS <- glm.nb(score.nbclaim, data=db.train)
```

Results

Comparison

	Poisson	NB2
(Intercept)	-1.5186	-1.5173
car_colorRed	-0.0215	-0.0212
need_glassesYes	0.1213	0.1216
territorySuburban	-0.0792	-0.0789
territoryUrban	-0.1242	-0.1245
languageFrench	0.2584	0.2586
foodVegan	-0.1015	-0.1027
foodVegetarian	-0.1963	-0.1971
\`$\alpha$\`	NA	2.1244

Prediction quality on the test set

Even if it is not the objective of this presentation, we can compare the fit and the prediction quality of the Poisson and the Negative Binomial 2. For the training set, the loglikelihood is used and a logarithmic score (LS) is used on the test set:

$$LS = -\sum_{i=1}^{m} \log(\Pr(N_i = n_i | \boldsymbol{X}_i))$$

Computation

```
db.test$pred <- predict(Poisson, newdata=db.test, type="response")
logs.Poisson <- -sum(dpois(db.test$NbClaims, db.test$pred, log=TRUE))
ll.Poisson <- logLik(Poisson)

db.test$pred <- predict(nb2.MASS, newdata=db.test, type="response")
alpha <- 1/nb2.MASS$theta
tau <- 1/nb2.MASS$theta
ll <- lgamma(db.test$NbClaims + alpha) - lgamma(alpha) - lgamma(db.test$NbClaims+1) + alpha*log(tau
logs.NB2 <- -sum(ll)
ll.NB2 <- logLik(nb2.MASS)
```

Comparison of models

Results

	Log-likelihood (Train)	Logarithmic Score (Test)
Poisson	-44,858.44	19,319.22
Negative Binomial 2	-44,718.43	19,549.53

Predictive ratemaking

Distribution of the heterogeneity

Based on the baysian model, by fitting a Negative Binomial distribution (NB2) on claim counts data, we know that the heterogeneity of our portfolio Θ is following a gamma($\alpha = 1.6654$, $\tau = \alpha = 1.6654$).

Predictive premiums

The predictive premium of an insured with $n_{i,\bullet} = \sum_{t=1}^{T-1} n_{i,t}$ past claims, and $\lambda_{i,\bullet} = \sum_{t=1}^{T-1} \lambda_{i,t}$ as the sum of past *a priori* premiums, is equal to:

$$E[N_{i,T}|n_{i,1}, \ldots, n_{i,T-1}, \boldsymbol{X}_{i,T}] = \lambda_{i,T} \frac{\alpha + \sum_{t=1}^{T-1} n_{i,t}}{\alpha + \sum_{t=1}^{T-1} \lambda_{i,t}} = \lambda_{i,T} \frac{1.6654 + n_{i,\bullet}}{1.6654 + \lambda_{i,\bullet}}$$

In STAM exam, we were able to analyse in details this equation.

Practical use

However, even if the Poisson-gamma model is theoretically correct, and even studied in the preliminary exams, this predictive rating approach is almost never used in practice:

- There is not weight in $\sum_{t=1}^{T-1} n_{i,t}$. That means that a claim from 10 or 20 years ago will have the same impact of the premium that an accident that was claimed last year;
- The value of $\sum_{t=1}^{T-1} \lambda_{i,t}$ depend on the estimated values of β, and should then be computed each year. That means that insurers should keep all past covariates $\boldsymbol{X}_{i,t}$, from $t = 1, \ldots, T$ of all their insureds. For new insureds, this is even more complicated.

Bonus-malus scales

To compute predictive premiums, actuaries have created **Bonus-Malus Scales** (BMS) models. BMS are class systems where the insured's level ℓ increases or decreases only by the number of claims.

Structure of the BMS (example with 6 levels)

Level	Relativities
6	1.296
5	1.197
4	1.141
3	1.017
2	0.986
1	0.879

Transition rules

A BMS is defined by its number of levels and by its transition rules. If we suppose a BMS with 6 levels (from 1 to 6), we can than create a BMS having the following rules:

- A new insured has an entry level 1;
- The BMS level of an insured without claim will be lowered by 1 (-1);
- The BMS level of an insured will increase by the number of claims times 2 (+2).

We then can summarize the transition rule system as:

	Level at time t+1, if x claims			
Starting Level (time t)	**x=0**	**x=1**	**x=2**	**x>3**
1	1	3	4	6
2	1	3	5	6
3	2	4	6	6
4	3	5	6	6
5	4	6	6	6
6	5	6	6	6

Transition matrix

This means that for a specific distribution, for example a Poisson or a NB2 distribution, it becomes possible to construct the transition matrix from time t to time $t+1$.

	Level at time t+1					
Level at time t	**1**	**2**	**3**	**4**	**5**	**6**
1	Pr(N=0)	0	Pr(N=1)	0	Pr(N=2)	Pr(N>2)
2	Pr(N=0)	0	0	Pr(N=1)	0	Pr(N>1)
3	0	Pr(N=0)	0	0	Pr(N=1)	Pr(N>1)
4	0	0	Pr(N=0)	0	0	Pr(N>0)
5	0	0	0	Pr(N=0)	0	Pr(N>0)
6	0	0	0	0	Pr(N=0)	Pr(N>0)

Calibration of the BMS

Covariates and risk characteristics

We can use the dataset of this example. We suppose an insured with some specific risk characteristics, for example, the insured might drive a black car, be an English speaker, be vegan, etc.

We will suppose a Poisson distribution, this insured has the following *a priori* premium: $\lambda = 0.2532$.

Heterogeneity

For simplicity, we will suppose two kinds of drivers, meaning the following distribution for Θ:

	Proportion	`$\theta$`	`$\lambda\theta$`
Good driver	66.7%	0.75	0.1899
Bad driver	33.3%	1.50	0.3798

Transition matrix for each type of drivers

With the mean parameter of the Poisson distribution, it becomes possible to compute the transition matrix for both insured:

Good driver (0.1899)

82.7%	0.0%	15.7%	0.0%	1.5%	0.1%
82.7%	0.0%	0.0%	15.7%	0.0%	1.6%
0.0%	82.7%	0.0%	0.0%	15.7%	1.6%
0.0%	0.0%	82.7%	0.0%	0.0%	17.3%
0.0%	0.0%	0.0%	82.7%	0.0%	17.3%
0.0%	0.0%	0.0%	0.0%	82.7%	17.3%

Bad driver (0.3798)

68.4%	0.0%	26.0%	0.0%	4.9%	0.7%
68.4%	0.0%	0.0%	26.0%	0.0%	5.6%
0.0%	68.4%	0.0%	0.0%	26.0%	5.6%
0.0%	0.0%	68.4%	0.0%	0.0%	31.6%
0.0%	0.0%	0.0%	68.4%	0.0%	31.6%
0.0%	0.0%	0.0%	0.0%	68.4%	31.6%

Stationnary matrix for each type of drivers

The long-term distribution of insured within the levels of the BMS can also be computed, where the initial BMS level at time $t = 1$ does not have any impact.

Good driver (0.1899)

57.7%	12.1%	14.6%	6.7%	5.8%	3.2%
57.7%	12.1%	14.6%	6.7%	5.8%	3.2%
57.7%	12.1%	14.6%	6.7%	5.8%	3.2%
57.7%	12.1%	14.6%	6.7%	5.8%	3.2%
57.7%	12.1%	14.6%	6.7%	5.8%	3.2%
57.7%	12.1%	14.6%	6.7%	5.8%	3.2%

Bad driver (0.3798)

24.0%	11.1%	16.2%	14.6%	17.1%	17.1%
24.0%	11.1%	16.2%	14.6%	17.1%	17.1%
24.0%	11.1%	16.2%	14.6%	17.1%	17.1%
24.0%	11.1%	16.2%	14.6%	17.1%	17.1%
24.0%	11.1%	16.2%	14.6%	17.1%	17.1%
24.0%	11.1%	16.2%	14.6%	17.1%	17.1%

BMS Relativities

Posterior distribution

Knowing the *a priori* distribution of the heterogeneity Θ, and the stationary distribution within the levels of the BMS for all types of drivers, we can compute the posterior distribution of Θ, conditional on the level ℓ:

$$\begin{aligned}\Pr[\text{Good driver}|L=\ell] &= \frac{\Pr[L=\ell|\text{Good driver}]\Pr[\text{Good driver}]}{\Pr[L=\ell|\text{Good driver}]\Pr[\text{Good driver}]+\Pr[L=\ell|\text{Bad driver}]\Pr[\text{Bad driver}]}\\ &= 1-\Pr[\text{Bad driver}|L=\ell]\end{aligned}$$

Conditional expectation

With the posterior distribution of Θ, the BMS relativity of each level can be computed with the conditional expectation:

$$r_\ell = E[\Theta|L=\ell] = 0.75\times\Pr[\text{Good driver}|L=\ell]+1.50\times\Pr[\text{Bad driver}|L=\ell]$$

Results

Simple calculations with our numerical example lead to:

	BMS Level					
	1	**2**	**3**	**4**	**5**	**6**
Good driver	82.8%	68.6%	64.3%	47.9%	40.4%	27.2%
Bad driver	17.2%	31.4%	35.7%	52.1%	59.6%	72.8%
BMS relativity	87.9%	98.6%	101.7%	114.1%	119.7%	129.6%

Computation of relativities

The results can be generalized with continuous heterogeneity, where it can be shown that BMS relativities are computed using:

$$r_\ell = \frac{\int_0^\infty \theta \Pi_\ell(\lambda\theta) g(\theta) d\theta}{\int_0^\infty \Pi_\ell(\lambda\theta) g(\theta) d\theta} \quad \text{, for } \ell = 1, \ldots, \ell_{max}.$$

where:

- $g(\theta)$ is the prior heterogeneity distribution;
- $\Pi_\ell(\lambda\theta)$ is the ℓ line component of $\Pi(\lambda\theta)$, the stationary distribution of insured of mean $\lambda\theta$.

Summary of the BMS

Characteristics of the BMS

To calculate the relativities, the actuary must select the characteristics of the BMS. Different choices will lead to different values of relativities $r_\ell, \ell = 1, \ldots, \ell_{max}$:

- The maximum number of levels ℓ_{max} of the BMS;
- The value of the penalty for each claim, i.e. the penalty structure of the BMS (ex: -1/+2);
- Other transition rules (for example: 3-5 years without claim automatically gives the largest discount);
- The entry level for new insureds.

Selection of the best BMS

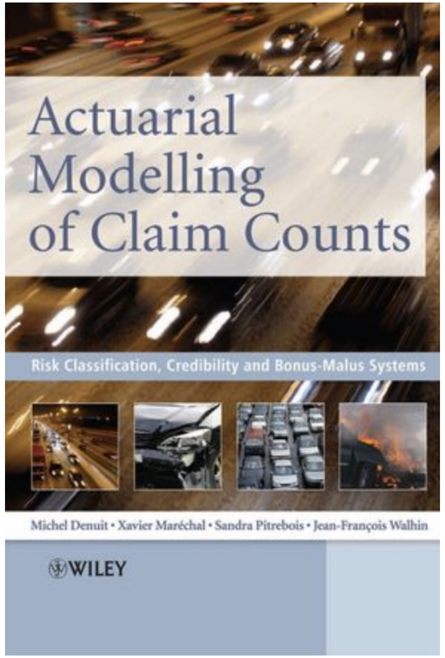
A variety of methods has been developed in the scientific literature to select the best BMS:

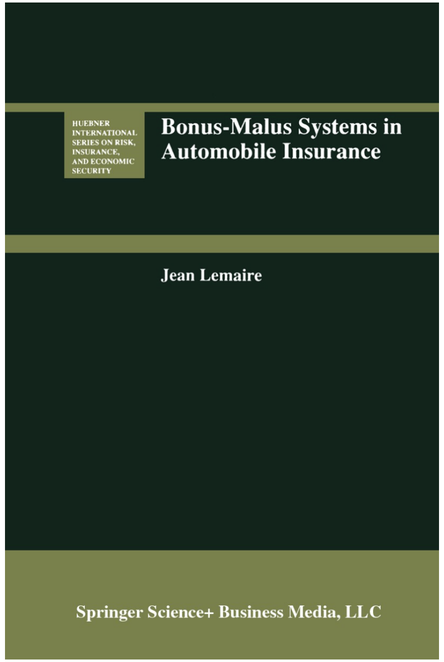
- The coefficient of variation;
- The mean-square error of prediction;
- The elasticity of the BMS.

For more details

Denuit, M., Maréchal, X., Pitrebois, S., & Walhin, J. F. (2007). *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons.

Lemaire, J. (1995). *Bonus-malus systems in automobile insurance (Vol. 19)*. Springer science & business media.





Part I - Ratemaking with Cross-Section Data

- Basic count distributions
- Credibility Models and Predictive Ratemaking
- Bonus-Malus Scales Models

Part II - Ratemaking with Panel Data

- Families of Count Distributions
- Observed Predictive Premiums
- Bonus-Malus Scales Models Revisited

Part III - Actual Challenges

- Entry levels and new insureds
- Penalties and *a priori* risks

Actual insurance database

Insureds (or even vehicles) are observed over time. Here, the dataframe *df2.Rda* contains 25,078 vehicles each observed for 5 years.

policy_no	veh.num	renewal_date	start_date	end_date	risk_expo	freq_payment	year_veh	sex	year_b
6000274	1	2011-02-03	2010-02-03	2011-02-02	1.0	12	2000	F	
6000274	1	2012-02-03	2011-02-03	2012-02-02	1.0	12	2000	F	
6000274	1	2013-02-03	2012-02-03	2013-02-02	1.0	12	2000	F	
6000274	1	2014-02-03	2013-02-03	2014-02-02	1.0	12	2000	F	
6000274	1	2015-02-03	2014-02-03	2014-12-28	0.9	12	2000	F	
6000517	1	2011-04-11	2010-04-11	2011-04-10	1.0	12	2006	M	

Claim count for panel data

Families of models

We have to suppose a form of dependance between all contracts of the same insured/vehicle. For count distributions, panel data modeling admits 3 families (see Molenberghs & Verbeke, 2005):

- Transition models (for example: time series for count data);
- Marginal approach (for example: Generalized Estimating Equations - GEE);
- Conditional approach with random effects.

Conditional approach

General form

In actuarial science, the conditional approach is the most popular approach. It that can be seen as a generalization of the heterogeneity approach seen earlier:

$$\Pr[N_{i,1} = n_{i,1}, \ldots, N_{i,T} = n_{i,T}] = \int_{D_\Theta} \left(\prod_{t=1}^{T} \Pr[N_{i,t} = n_{i,t} | \theta] \right) g(\theta) d\theta.$$

The Poisson-gamma model revisited

A conditional Poisson distribution, with gamma random effects leads to the multivariate negative binomial distribution (MVNB), a generalization of the NB2 distribution:

$$\Pr[N_{i,1} = n_{i,1}, \ldots, N_{i,T} = n_{i,T}] = \left(\prod_{t=1}^{T} \frac{(\lambda_{i,t})^{n_{i,t}}}{n_{i,t}!} \right) \frac{\Gamma(\sum_{t=1}^{T} n_{i,t} + \alpha)}{\Gamma(\alpha)} \left(\frac{\alpha}{\sum_{i=1}^{T} \lambda_{i,t} + \alpha} \right)^{\alpha} \left(\sum_{i=1}^{T} \lambda_{i,t} + \alpha \right)^{-\sum_{t=1}^{T} n_{i,t}}.$$

Posterior ratemaking

As for cross-section data, we are interested to compute the predictive premium. The predictive premium of an insured with $\sum_{t=1}^{T-1} n_{i,t}$ past claims, and $\sum_{t=1}^{T-1} \lambda_{i,t}$ as the sum of past *a priori* premiums, is equal to:

$$E[N_{i,T}|n_{i,1},\ldots,n_{i,T-1},\boldsymbol{X}_{i,T}] = \lambda_{i,T}\frac{\alpha + \sum_{t=1}^{T-1} n_{i,t}}{\alpha + \sum_{t=1}^{T-1} \lambda_{i,t}}.$$

This result is similar to what we obtained for the predictive premium of a NB2 distribution.

Same problems

For the same reasons why the NB2 was not used in practice (no weight in $\sum_{t=1}^{T-1} n_{i,t}$, and the need to use past $\lambda_{i,t}$, for example), another approach has to be used in practice.

As before, by knowing the distribution of the random effects Θ, the Bonus-Malus Scale models can be an interesting solution...

Predictive distribution

Even if the situation is similar, cross-section data and panel data models are not the same.

Joint distribution

The joint distribution for all the contracts of the same insured can be rewritten as:

$$\mathbf{Pr}[N_{i,1} = n_{i,1}, \ldots, N_{i,T} = n_{i,T}] = \mathbf{Pr}[N_{i,1} = n_{i,1}] \times \mathbf{Pr}[N_{i,2} = n_{i,2} | n_{i,1}] \times \ldots \times \mathbf{Pr}[N_{i,T} = n_{i,T} | n_{i,1}, \ldots, n_{i,T-1}]$$

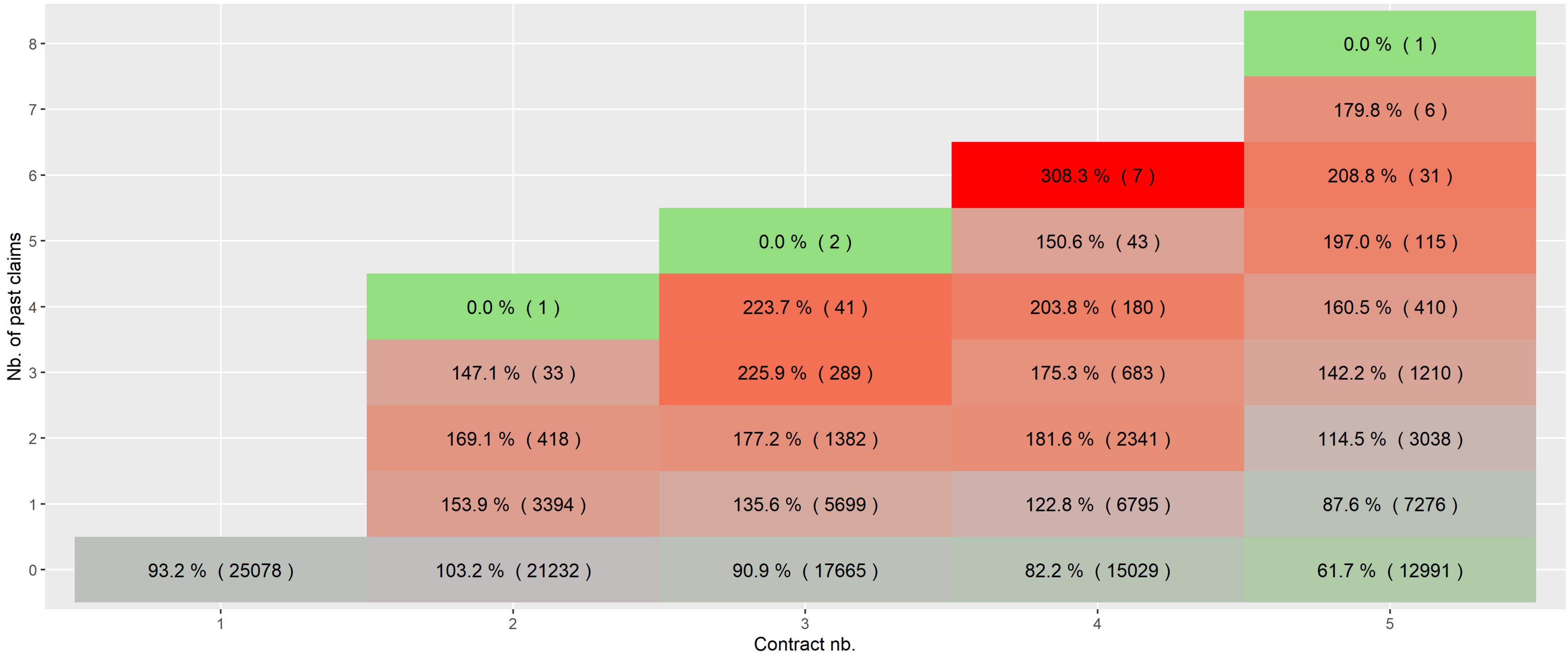
That means that the distribution the predictive distribution $\mathbf{Pr}[N_{i,t} = n_{i,t} | n_{i,1}, \ldots, n_{i,t-1}]$ is already used in the modeling and thus, a predictive premium is already computed in the underlying model.

Assumption

We do not need to only rely on the assumption of the constant heterogeneity term to ccompute the predictive premium (even more for large longitudinal dataset - as for the farm insurance dataset used in the published papers).

Empirical analysis

We can indeed verify empirically the values of the predictive premiums (as a percentage of the average frequency).



Bonus-Malus scales

BMS are still interesting for actuaries and insurers:

- Advanced panel data models based on random effects, hierachical copulas, etc. cannot be easily used for ratemaking in practice;
- The penalty structure of BMS are well-known by many insurers, brokers, regulators and insureds, and easy to explain/understand;
- BMS allow complex penalty structure that might be difficult to implement with classic statistical models:
    - Fast-track for forgiveness (ex: for example: 3-5 years without claim automatically gives the largest discount);
    - Multi-vehicules penalty structure;
    - Multi-products penalty structure;
    - etc.
- There is a large scientific literature on BMS that can be used.

The challenge with BMS

The problem is not on the BMS itself, but on how BMS can now be estimated with the current longitudinal/hierarchical data of insurers:

- We do not have to rely on the long-term behaviour of the insureds, based on the heterogeneity distribution;
- BMS relativities have to be estimated directly with the data;
- A direct comparison between BMS premiums and data can be done.

Boucher & Inoussa (2014) were the first to be interested in adapting the Bonus-Malus Scales approach to the new databases of insurers.

In the following slides, a link between the classic GLM approach and the BMS is proposed to better understand how actuaries can estimated all the BMS parameters.

Using past informations as covariates

Instead of using the bayesian approach, with an unknown risk profile that be updated after each contract, many insurers directly include past claims information as covariate in the μ mean parameter of the count distribution.

For example, for an insured with T years of experience, some actuaries use:

$$\mu_{i,T} = \exp\Big(X'_{i,T}\beta + \gamma_1 n_{i,T-1} + \gamma_2 n_{i,T-2} + \ldots + \gamma_1 n_{i,1}\Big)$$

Sum of past claims

This means a large amount of parameters $\gamma_1, \ldots, \gamma_a$. (One of) the purpose of statistics is to summarize information. One classic approach is instead to use a summary of past claims. For example, we can use:

$$\mu_{i,t} = \exp(X'_{i,t}\beta + \gamma n_{i,\bullet})$$

where $n_{i,\bullet}$ is the number of all past claims for insured i.

New insureds and insureds with experience

The problem with the last approach is that we cannot differentiate new insureds from insureds with many years of experience: both types of insureds have $n_{i,\bullet} = 0$. Instead, we should use:

$$\mu_{i,T} = \exp(\boldsymbol{X}_{i,T}'\beta + \gamma_1 \kappa_{i,\bullet} + \gamma_2 n_{i,\bullet})$$

where, for insured i:

- $n_{i,\bullet} = \sum_{t=1}^{T-1} n_{i,t}$ is the number of all past claims;
- $\kappa_{i,\bullet} = \sum_{t=1}^{T-1} I(n_{i,t} = 0)$ is the sum of policy periods without claims.

This allows us to differentiate new insureds from insureds with many years of experience.

This model is called the **Kappa-N model**.

Fitting the Kappa-N model

The Kappa-N model can be used with any distribution. The Poisson Kappa-N and the NB2 Kappa-N are presented below:

Add past information

```
data <- df2 %>%
  mutate(ind.0 = (NbClaims == 0)) %>%
  arrange(policy_no, veh.num, renewal_date) %>%
  group_by(policy_no, veh.num) %>%
  mutate(contract.no = row_number(),
         past.n = cumsum(NbClaims)- NbClaims,
         past.kappa = cumsum(ind.0) - ind.0) %>%
  ungroup()
```

Split data

```
db.train <- data %>% filter(Type=='TRAIN')
db.test <- data %>% filter(Type=='TEST')
```

Poisson Kappa-N and NB2 Kappa-N models

Fitting models

```
score.nbclaim <- as.formula(NbClaims ~ car_color + need_glasses + territory + language + food + past
Poisson    <- glm(score.nbclaim, family=poisson(link=log), data=db.train)
nb2.MASS <- glm.nb(score.nbclaim, data=db.train)
```

Poisson Kappa-N and NB2 Kappa-N models

	Poisson	NB2(MASS)
(Intercept)	-1.5651	-1.5641
car_colorRed	-0.0251	-0.0257
need_glassesYes	0.1001	0.1006
territorySuburban	-0.0734	-0.0730
territoryUrban	-0.1174	-0.1180
languageFrench	0.2284	0.2290
foodVegan	-0.0047	-0.0062
foodVegetarian	-0.0272	-0.0283
past.n	0.1976	0.2004
past.kappa	-0.1101	-0.1101
\`$\alpha$\`	NA	2.4866

Estimated parameters

The mean parameter of the Poisson and the NB2 distribution was defined as:

$$\lambda_{i,t} = \exp(X'_{i,t}\beta + \gamma_1\kappa_{i,\bullet} + \gamma_2 n_{i,\bullet})$$

Results for the fictive car insurance dataset:

For the Poisson distribution, we obtained:

- $\widehat{\gamma}_1$ = 0.1976;
- $\widehat{\gamma}_2$ = -0.1101.

Results for the Farm dataset (in the published papers):

For the Poisson distribution, we obtained:

- $\widehat{\gamma}_1$ = 0.0935;
- $\widehat{\gamma}_2$ = -0.0238.

Rewriting the model

We can rewrite the Kappa-N model with the following steps:

1- Instead of using $\kappa_{i,\bullet}$ as a covariate, we used a slightly modified transformation: $100 - \kappa_{i,\bullet}$.

- The negative value in front of $\kappa_{i,\bullet}$ helps to understand that high values of contracts without claim should decrease the premium;
- The value of 100 will be used as the entry level for insureds without experience.

$$\lambda_{i,t} = \exp\Big(\boldsymbol{X}'_{i,t}\beta^* + \gamma_1(100 - \kappa_{i,\bullet}) + \gamma_2 n_{i,\bullet}\Big)$$

2- We factor out the parameter γ_1 to obtain:

$$\lambda_{i,t} = \exp\left(\boldsymbol{X}'_{i,t}\beta^* + \gamma_1\left(100 - \kappa_{i,\bullet} + \frac{\gamma_2}{\gamma_1} n_{i,\bullet}\right)\right) = \exp\Big(\boldsymbol{X}'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\Big)$$

where the parameter $\ell_{i,t} = \left(100 - \kappa_{i,\bullet} + \frac{\gamma_2}{\gamma_1} n_{i,\bullet}\right)$ can be seen as a **claim score**.

Penalty structure

With the mean parameter:

$$\lambda_{i,t} = \exp\left(\boldsymbol{X}'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\right), \text{ where } \ell_{i,t} = \left(100 - \kappa_{i,\bullet} + \frac{\gamma_2}{\gamma_1} n_{i,\bullet}\right)$$

Details

1) For new insured, without insured experience, we have $n_{i,\bullet} = 0$, and $\kappa_{i,\bullet} = 0$, which means an initial claim score of 100.

2) Each year without claim decrease his claim score by 1.

3) Each claim increases the claim score by $\Psi = \frac{\widehat{\gamma_2}}{\widehat{\gamma_1}} = 3.93$, called the *jump-parameter*:

- One claim equals $\approx \Psi$ years without claims;
- $\Psi = \frac{0.1976}{0.1101} = 1.79$ for the fictive car insurance dataset.

Penalty structure (2)

With the mean parameter:

$$\lambda_{i,t} = \exp\left(\boldsymbol{X}'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\right), \text{ where } \ell_{i,t} = \left(100 - \kappa_{i,\bullet} + \frac{\gamma_2}{\gamma_1} n_{i,\bullet}\right)$$

Details

4) The penalty for a claim is equal to:

- $\exp(0.1101 \times 1.79) - 1 = 21.78\%$ for the fictive car insurance dataset.

5) Each year without claim decreases the premium by:

- $1 - \exp(-0.1101) = 10.4\%$ for the fictive car insurance dataset.

Problem of the Kappa-N models

One obvious problem with the Kappa-N models is the possible extreme values of $\ell_{i,t}$. For the fictive dataset, we have:

1) Maximum value of $n_{i,\bullet}$: 8.

2) Maximum value for $\ell_{i,t}$: 114.4

- Results in a premium almost 4 times higher than the premium for a new insured;
- It would take 14 consecutive years without claim for this insured to have the same premium as a new insured.

3) Minimum value for $\ell_{i,t}$: 96.

- Discount of 35%.

A possible solution

One solution could be to limit the value of $\ell_{i,t}$ in the modeling.

For example, we can limit $\ell_{i,T}$ to be between 95 and 110, meaning $\ell_{min} = 98$ and $\ell_{max} = 105$:

- Instead of $\ell_{i,T} = 114.4$, an insured would have $\ell_{i,T} = 105$
  - ...but it would however still need him 14 consecutive years without claim to reach level 100!
- Instead of $\ell_{i,T} = 96$, the insured without claim would have $\ell_{i,T} = 98$
  - ...but it means that he could claim without having any surcharge!

A better solution

Instead of:

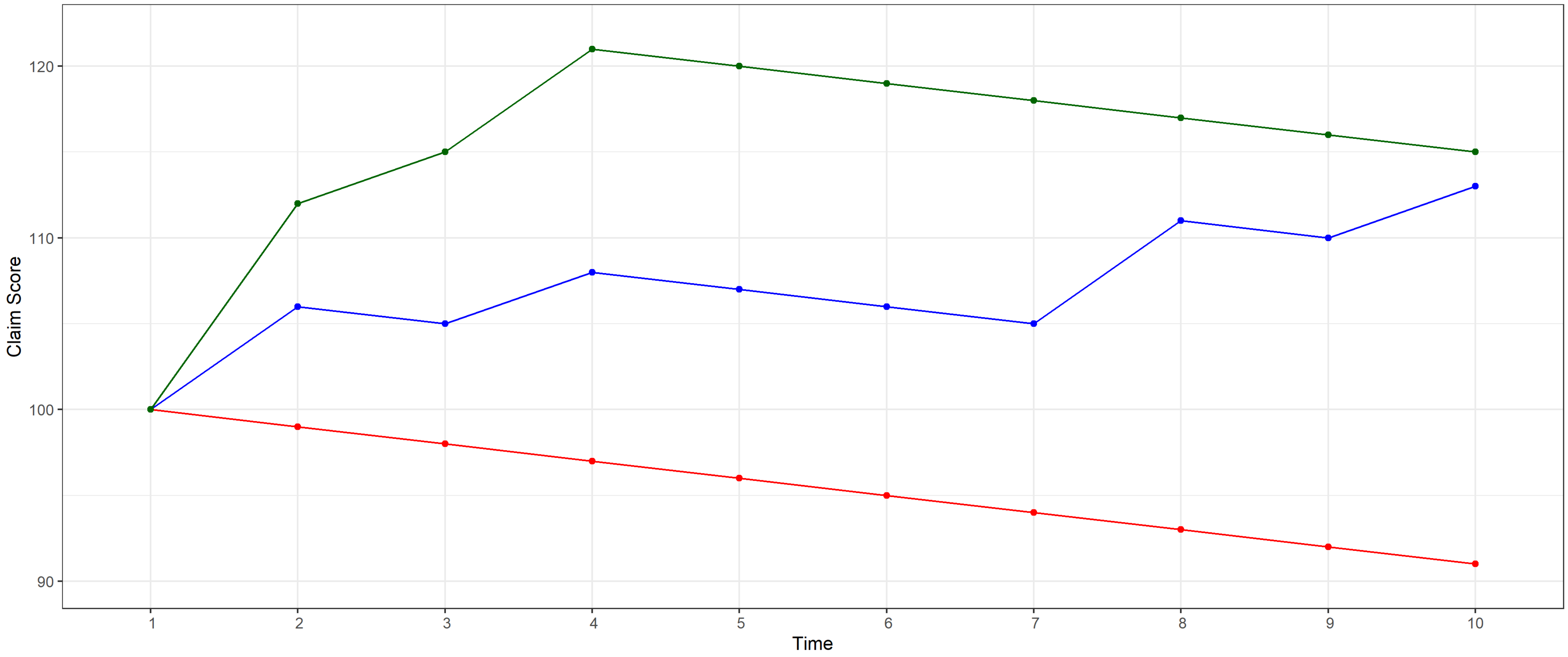
- limiting the claim score $\ell_{i,t}$ for **the current contract** $t = T$,
- we could limit the value of the claim score $\ell_{i,t}$ but for **all past contracts** $t = 1, \ldots, T$.

Simple illustration

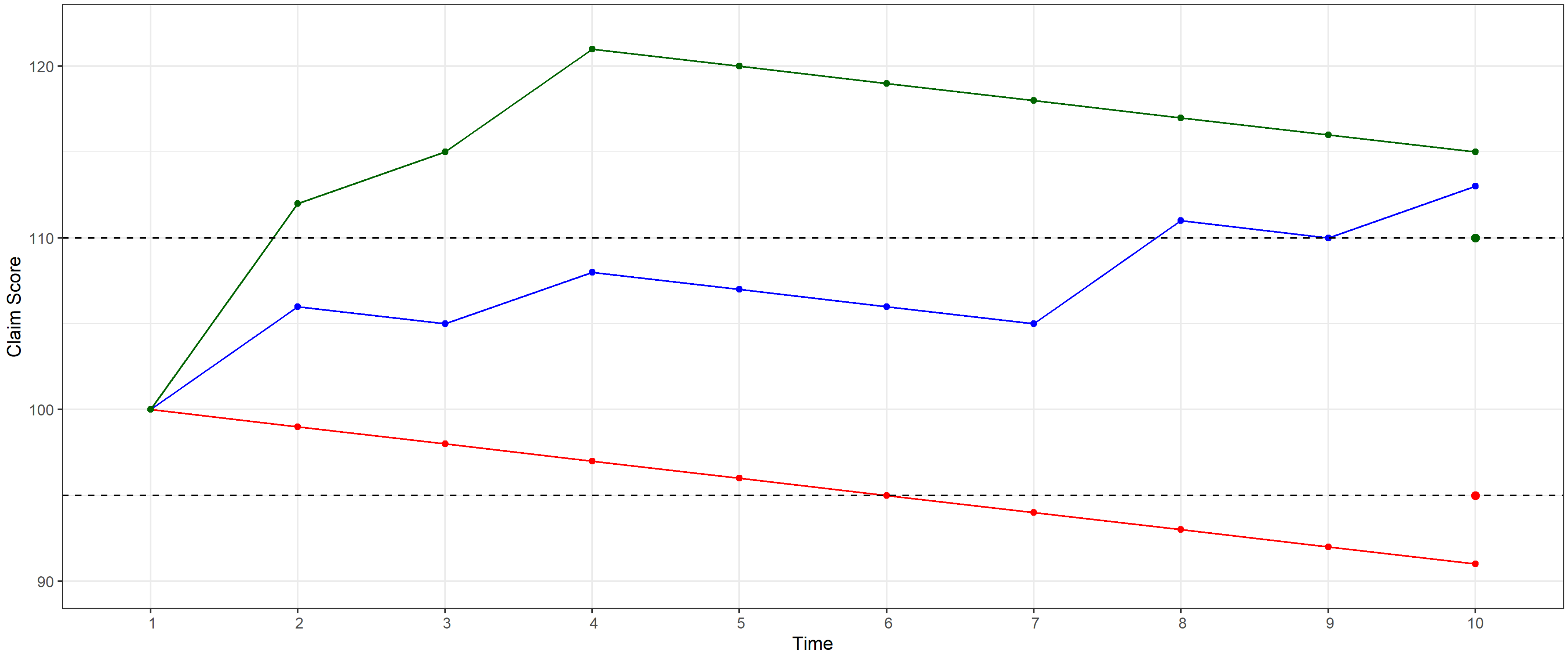
What happens to the claim score with a jump parameter $\Psi = 3$.

	Number of claims at time (t)									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Insured 1	0	0	0	0	0	0	0	0	0	0
Insured 2	2	0	1	0	0	0	2	0	1	0
Insured 3	4	1	2	0	0	0	0	0	0	0

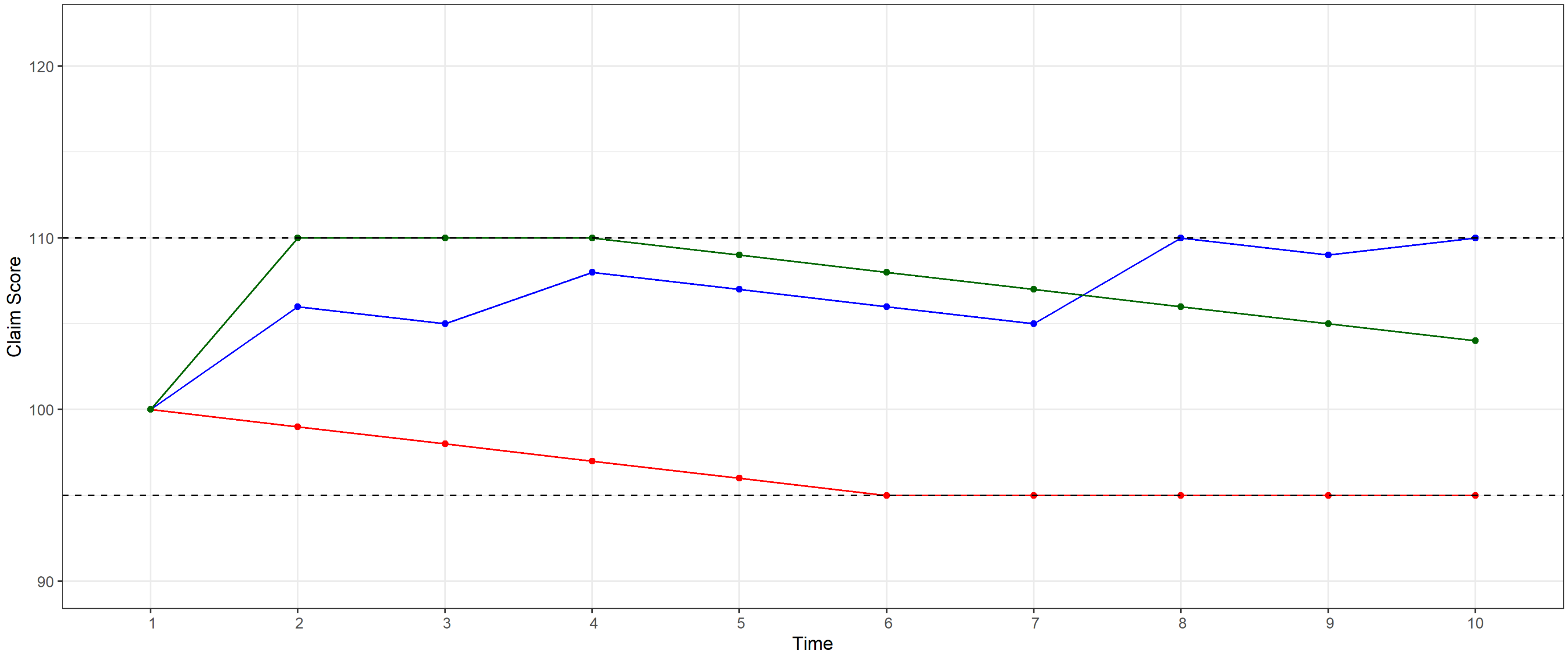
How to limit the claim score



How to limit the claim score



How to limit the claim score



Kappa-N model becomes a BMS model

By limiting the claim score for all past contracts, the Kappa-N model becomes a Bonus-Malus Scale Model. The claim score of insured i at time T, $\ell_{i,T}$, can now be seen as a BMS level.

A BMS without limits $\ell_{min} \to -\infty$ and $\ell_{max} \to \infty$ is a Kappa-N model.

Joint Distribution of all contracts

The joint distribution can now be expressed as the product of simple count distributions (with mean that depends on the Bonus-Malus level):

$$\mathbf{Pr}[N_{i,1} = n_{i,1}, \ldots, N_{i,T} = n_{i,T}] = \mathbf{Pr}[N_{i,1} = n_{i,1} | \ell_{i,1}] \times \mathbf{Pr}[N_{i,2} = n_{i,2} | \ell_{i,2}] \times \ldots \times \mathbf{Pr}[N_{i,T} = n_{i,T} | \ell_{i,T}]$$

where the markovian property of the BMS level can be used with:

$$\ell_{i,t} = \min(\max(\ell_{i,t-1} - I(n_{i,t-1} = 0) + \Psi \times n_{i,t-1}, \ell_{min}), \ell_{max})$$

Impact of the structural parameters of the BMS

To summarize, the BMS level $\ell_{i,t}$ depends on:

1- The jump parameter Ψ;
2- The minimum limit ℓ_{min};
3- The maximum limit ℓ_{max}.

Changing one of these three structural parameters will also changes the value of $\ell_{i,t}$, which means that the mean parameter of the count distribution of $N_{i,t}$ will change.

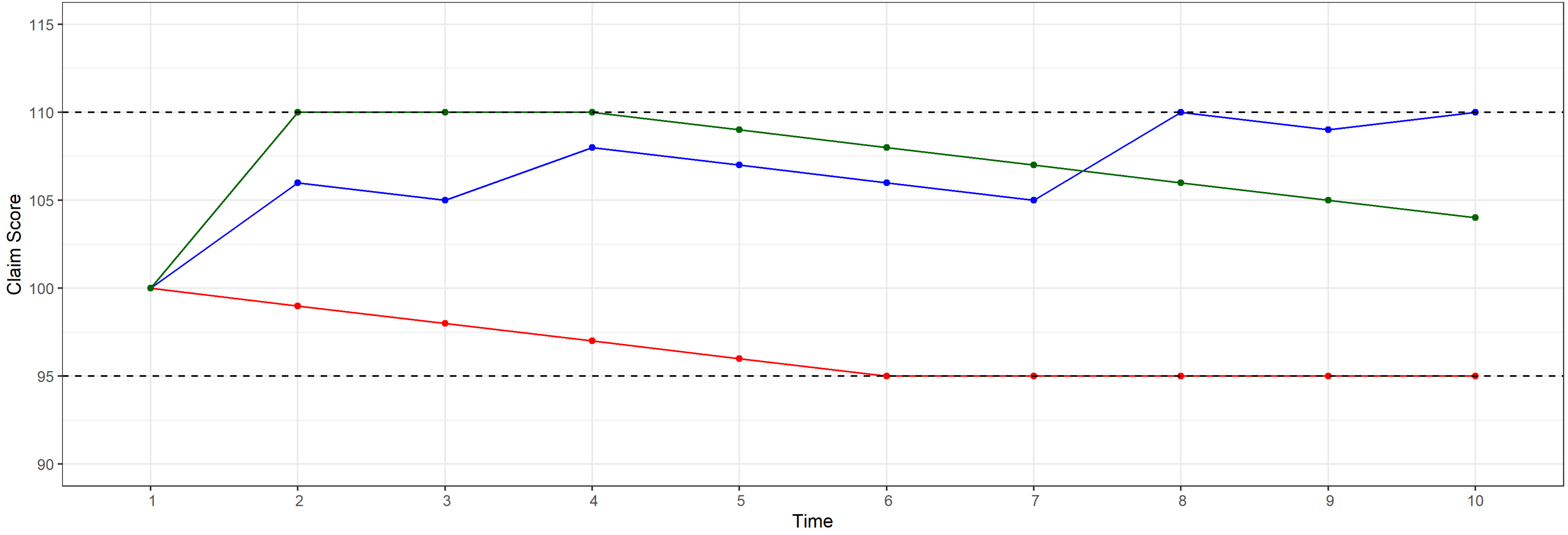
BMS level path

It is important to understand that for each combinaison of the structural parameters Ψ, ℓ_{min} and ℓ_{max}, the whole experience of each insured must be recomputed to obtain the correct BMS levels $\ell_{i,t}$.

Impact: example

Choice of structural parameters

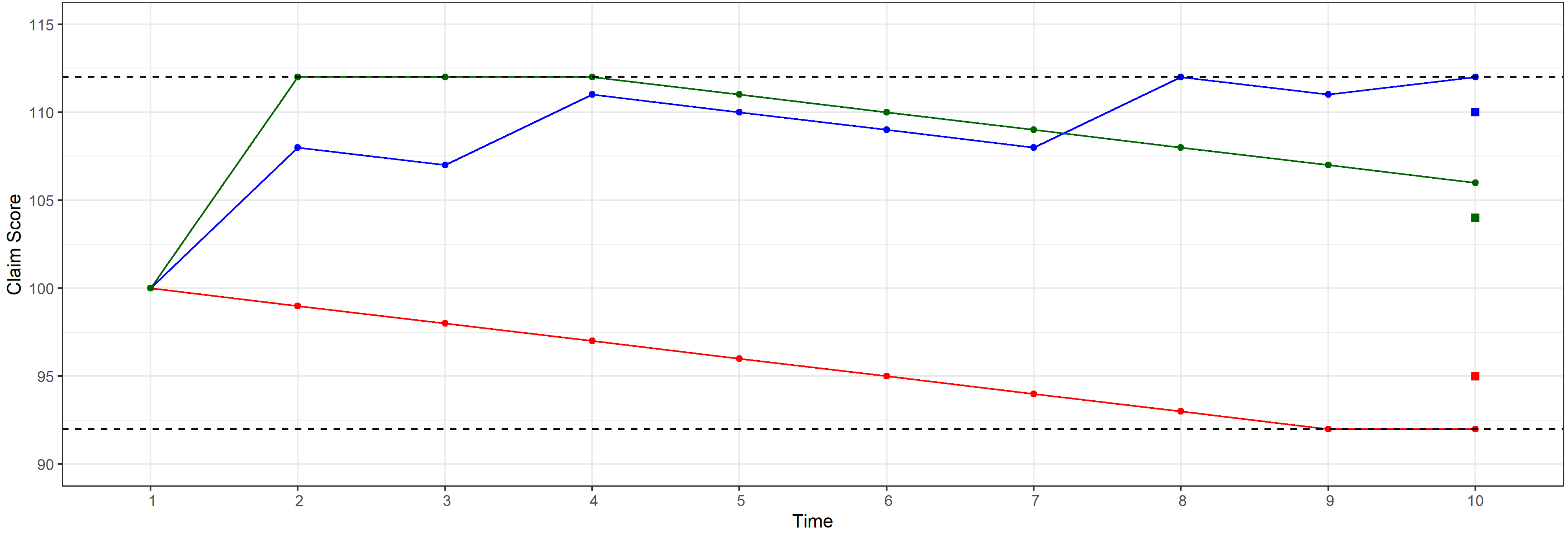
With $\Psi = 3, \ell_{min} = 95, \ell_{max} = 110$:



Impact: example

Choice of structural parameters

With $\Psi = 4, \ell_{min} = 92, \ell_{max} = 112$:



Parameters inference

Regresion parameters

For any insurance database, when the structural parameters are set, we can now compute the Bonus-Malus level of all contracts of each insured. We then apply a simple regression model with mean:

$$\lambda_{i,t} = \exp\Big(X'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\Big),$$

and estimate the parameters β, γ_1 and other parameters from the distribution (a dispersion parameter for example). The classic *GLM* package can be used for a Poisson, and the *MASS* package for the negative binomial.

Structural parameters

When the structural parameters Ψ, ℓ_{min} and ℓ_{max} are selected, it is easy to estimate the parameters of the BMS model.

But finding the best values of Ψ, ℓ_{min} and ℓ_{max} cannot be done directly. Even if we limit the strutural parameters to be integer, computing all possibilities might be too long.

For small dataset

For a small dataset such as the one used in this presentation, we can simply test all possibles values of the structural parameters.

```
Psi     <- seq(1, 10, length.out = 10)
ell.min <- seq(96, 99, length.out = 4)
ell.max <- seq(101, 120, length.out = 20)
grid <- expand.grid(ell.max = ell.max, ell.min = ell.min, Psi = Psi)
grid$llPoisson <- NA
grid$llNB2 <- NA

for(ii in 1:nrow(grid)){
  data         <- set.BMS_levels(ell.max=grid[ii,1], ell.min=grid[ii,2], Psi=grid[ii,3], db.train)
  PoissonBMS   <- glm(score.nbclaim, family=poisson(link=log), data=data)
  nb2BMS       <- glm.nb(score.nbclaim, data=data)
  grid[ii,4] <- logLik(PoissonBMS)
  grid[ii,5] <- logLik(nb2BMS)
}
print(grid[grid$llPoisson==max(grid$llPoisson),])
print(grid[grid$llNB2==max(grid$llNB2),])
```

The best BMS model, for the Poisson and the NB2 distributions, is $\ell_{max} = 104, \ell_{max} = 96, \Psi = 2$.

Proposed algorithm

For real insurance data, testing all possibilities is too long. A proposed iterative technique based on profile log-likelihood works as follow:

Initial step:

We set $\ell_{min}^{(0)} \to -\infty$ and $\ell_{max}^{(0)} \to \infty$ (this represents the Kappa-N model). We can then directly estimate a first estimate of the jump $\Psi^{(0)} = \Psi$.

For step $k = 1, \ldots$:

- With $\Psi = \Psi^{(k-1)}$ and $\ell_{min} = \ell_{min}^{(k-1)}$, we estimate all possible BMS models for any value of ℓ_{max}.
  - We choose $\ell_{max}^{(k)} = \ell_{max}$ from the best BMS model.
- With $\Psi = \Psi^{(k-1)}$ and $\ell_{max} = \ell_{max}^{(k)}$, we estimate all possible BMS models for any value of ℓ_{min}.
  - We choose $\ell_{min}^{(k)} = \ell_{min}$ from the best BMS model.
- With $\ell_{max} = \ell_{max}^{(k)}$ and $\ell_{min} = \ell_{min}^{(k)}$, we estimate all possible BMS models for any value of Ψ.
  - We choose $\Psi^{(k)} = \Psi$ from the best BMS model.

We repeat these steps until we reach convergence.

Results obtained with the fictive dataset (Poisson)

	Poisson	Poisson Kappa-N	Poisson BMS (104/96/+2)
(Intercept)	-1.5186	-1.5651	-13.3672
car_colorRed	-0.0215	-0.0251	-0.0217
need_glassesYes	0.1213	0.1001	0.0986
territorySuburban	-0.0792	-0.0734	-0.0745
territoryUrban	-0.1242	-0.1174	-0.1167
languageFrench	0.2584	0.2284	0.2275
foodVegan	-0.1015	-0.0047	-0.0070
foodVegetarian	-0.1963	-0.0272	-0.0268
\`$\gamma_1$\`	NA	0.1976	0.1180
\`$\gamma_2$\`	NA	-0.1101	NA

Analyzing the model

Expected number of claims

The mean of the BMS model can then be expressed as:

$$\lambda_{i,t} = \exp\Big(X'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\Big) = \exp\Big(X'_{i,t}\beta^* + \gamma_1 \ell_{i,t}\Big) = \Pi_{i,t} \times r(\ell_{i,t})$$

where:

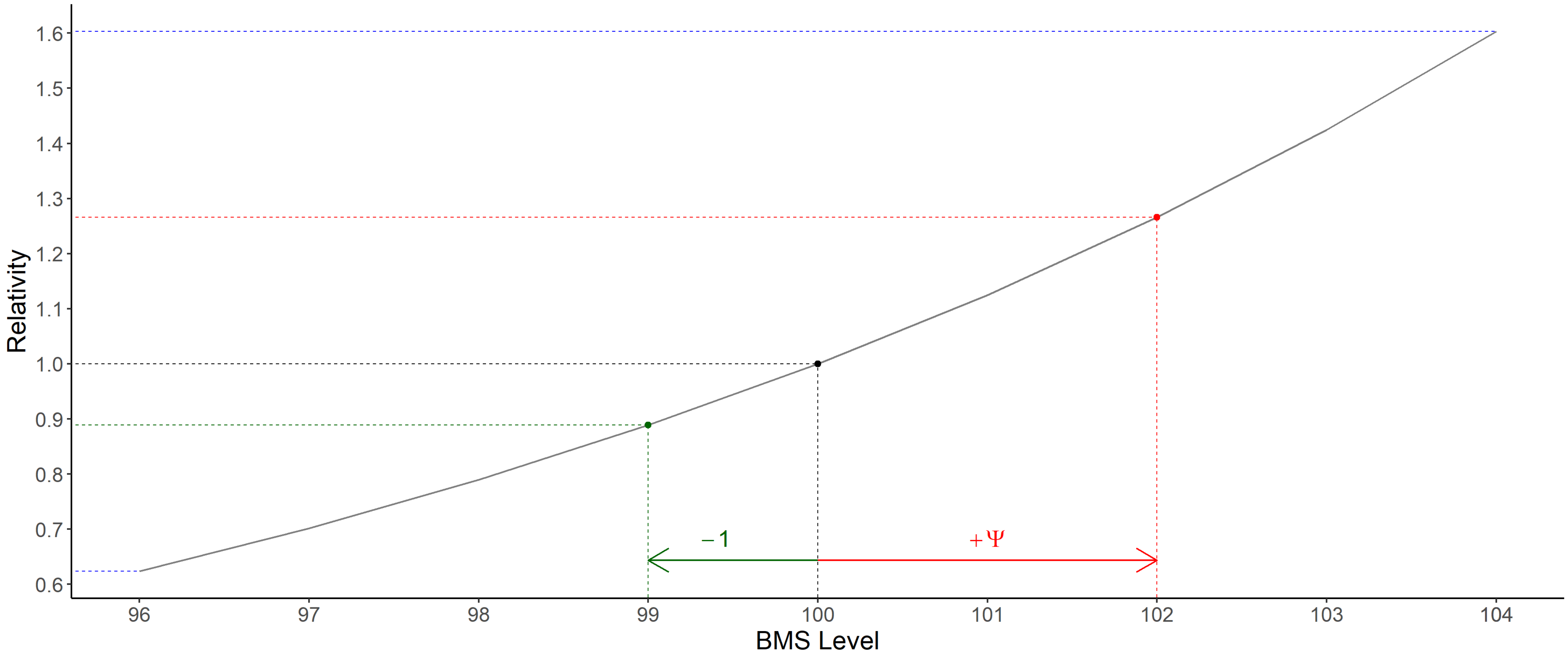
- $\Pi_{i,t} = \exp\Big(X'_{i,t}\beta^*\Big)$ is the base premium that depends on covariates for the contract t of insured i;
- $r(\ell_{i,t}) = \exp(\gamma_1 \ell_{i,t})$ is the BMS relativity that depends on the BMS level $\ell_{i,t}$ the contract t of insured i.

Prior and posterior ratemaking

As opposed to classic BMS calibration, all parameters of the mean are estimated simulatenously:

i) the β from the covariates (for the $\Pi_{i,t}$ component),
ii) the γ_1 for the predictive ratemaking (for the $r(\ell_{i,t})$ component).

Graph of BMS relativities



Fitting and predictive quality

Log-likelihood (Training set)

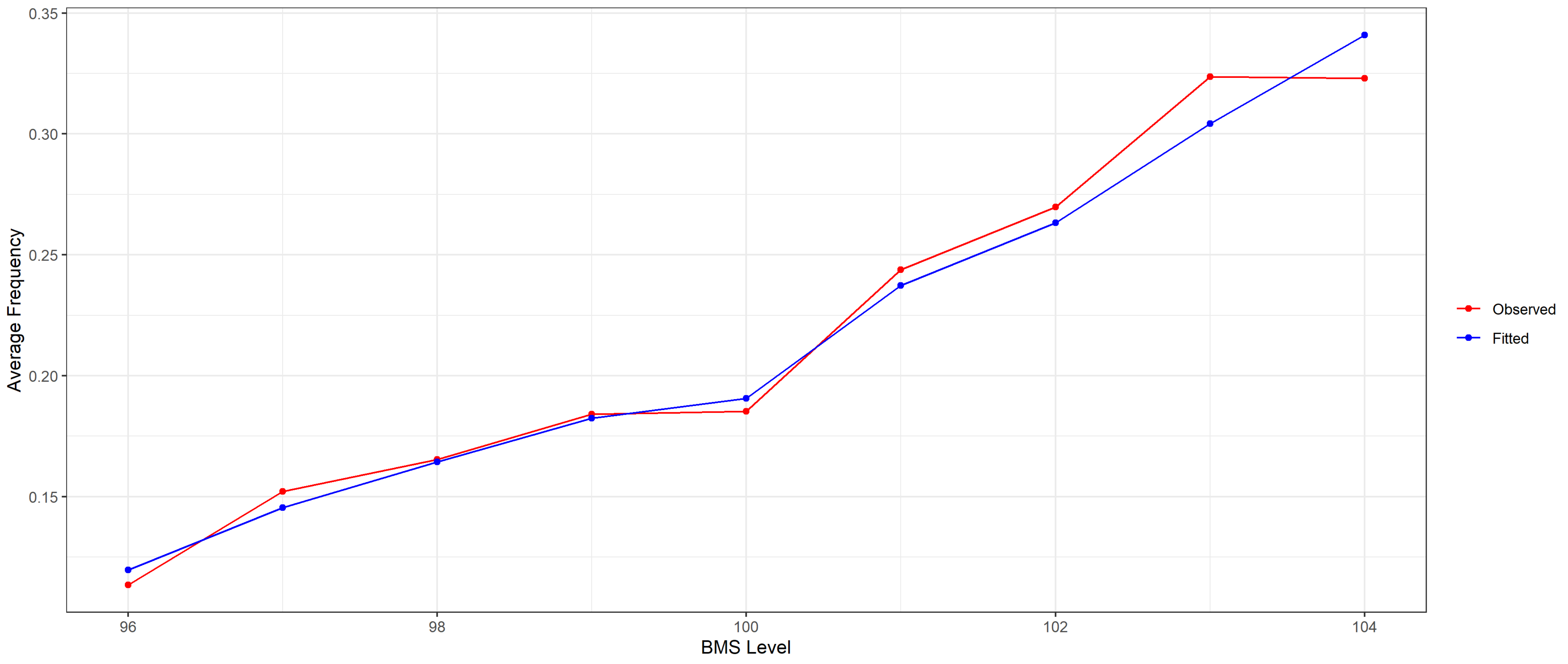
A correction, such as the AIC/BIC, must be applied for each model because they do not have he same number of parameters.

	Standard	Kappa-N	BMS
Poisson	-44,858.44	-44,461.88	-44,447.66
Negative Binomial 2	-44,718.43	-44,352.62	-44,340.49

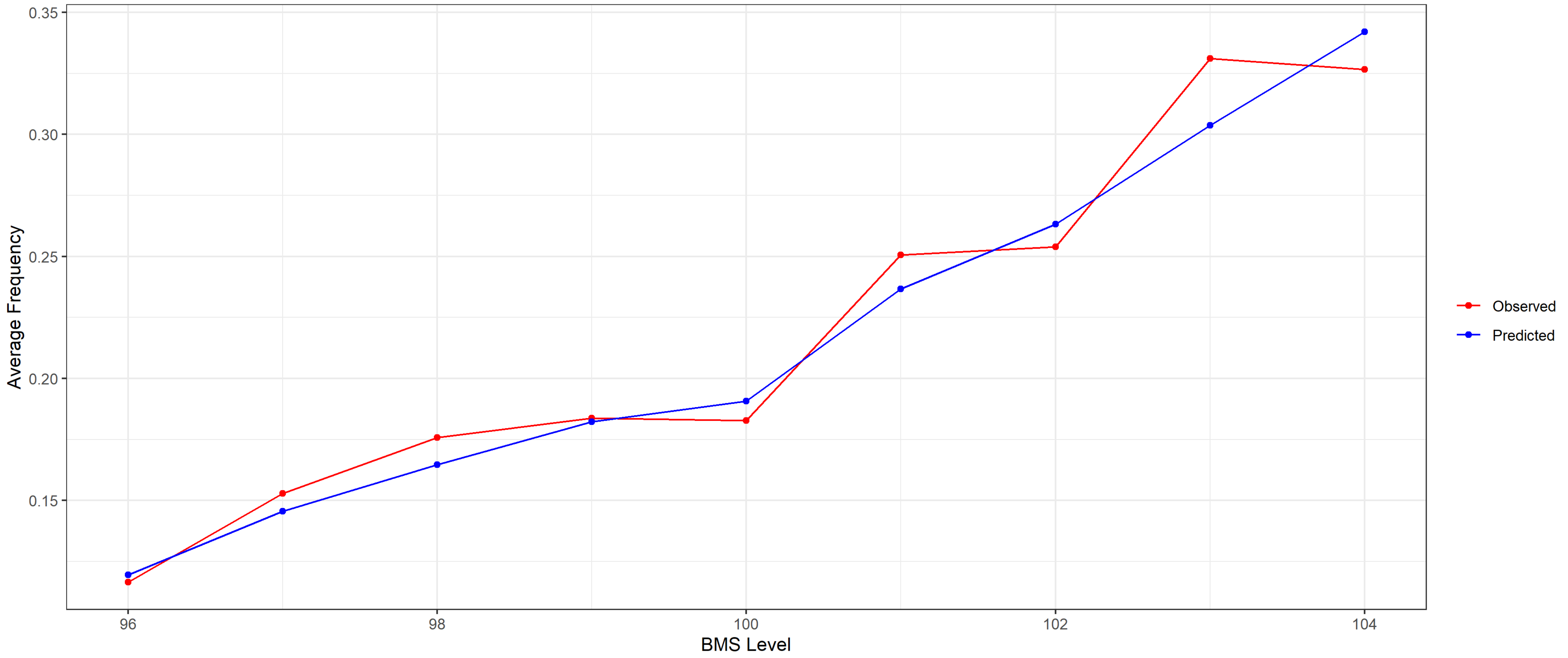
Logarithmic Score (Test set)

	Standard	Kappa-N	BMS
Poisson	19,319.22	19,169.51	19,160.35
Negative Binomial 2	19,549.53	19,539.31	19,542.47

Distribution over the BMS levels (training dataset)



Distribution over the BMS levels (test dataset)



Part I - Ratemaking with Cross-Section Data

- Basic count distributions
- Credibility Models and Predictive Ratemaking
- Bonus-Malus Scales Models

Part II - Ratemaking with Panel Data

- Families of Count Distributions
- Observed Predictive Premiums
- Bonus-Malus Scales Models Revisited

Part III - Actual Challenges

- Entry levels and new insureds
- Penalties and *a priori* risks

Reference

This part of the presentation is based on Section 4 of:

J.-P. Boucher (2022). Bonus-Malus Scale Models: Creating Artificial Past Claims History. *Annals of Actuarial Science*, 1-27.

Annals of Actuarial Science (2022), 1–27
doi:10.1017/S1748499522000100

Institute and Faculty of Actuaries

ORIGINAL RESEARCH PAPER

Bonus-Malus Scale models: creating artificial past claims history

Jean-Philippe Boucher

Chaire Co-operators en analyse des risques actuariels, Departement de mathematiques, Université du Québec à Montréal, Montreal, Canada
E-mail: boucher.jean-philippe@uqam.ca

(Received 04 January 2022; revised 09 June 2022; accepted 24 June 2022)

Abstract

In recent papers, Bonus-Malus Scales (BMS) estimated using data have been considered as an alternative to longitudinal data and hierarchical data approaches to model the dependence between different contracts for the same insured. Those papers, however, did not discuss in detail how to construct and understand BMS models, and many of the BMS's basic properties were not discussed. The first objective of this paper is to correct this situation by explaining the logic behind BMS models and by describing those properties. More particularly, we will explain how BMS models are linked with simple count regression models that have covariates associated with the past claims experience. This study could help actuaries to understand how and why they should use BMS models for experience rating. The second objective of this paper is to create artificial past claims history for each insured. This is done by combining recent panel data theory with BMS models. We show that this addition significantly improves the prediction capacity of the BMS and provides a temporary solution for insurers who do not have enough historical data. We apply the BMS model to real data from a major Canadian insurance company. Results are analysed deeply to identify specific aspects of the BMS model.

Joint distribution

We already mentioned that the joint distribution of all claim counts of each insured $i = 1, \ldots, m$ can be expressed as:

$$\Pr[N_{i,1} = n_{i,1}, \ldots, N_{i,T} = n_{i,T}] = \Pr[N_{i,1} = n_{i,1}|\ell_{i,1}] \times \Pr[N_{i,2} = n_{i,2}|\ell_{i,2}] \times \ldots \times \Pr[N_{i,T} = n_{i,T}|\ell_{i,T}]$$

where the markovian property of the BMS level can be used with:

$$\ell_{i,t} = \min(\max(\ell_{i,t-1} - I(n_{i,t-1} = 0) + \Psi \times n_{i,t-1}, \ell_{min}), \ell_{max})$$

New insureds

We have a problem with $\Pr[N_{i,1} = n_{i,1}|\ell_{i,1}]$: thos insureds at time $t = 1$ are not always new drivers, but often new insureds in the company or new insured in the database. For the first contract at $t = 1$, we do not have $\ell_{i,0}$ nor $n_{i,0}$.

The major problem with past claims rating refers to the availability of past information:

- Insurers are not able to obtain past information from other insurers;
- Insurers are also often unable to use information from their own old contracts (modification of their operating systems, when past databases are simply erased or are no longer useful, etc.).

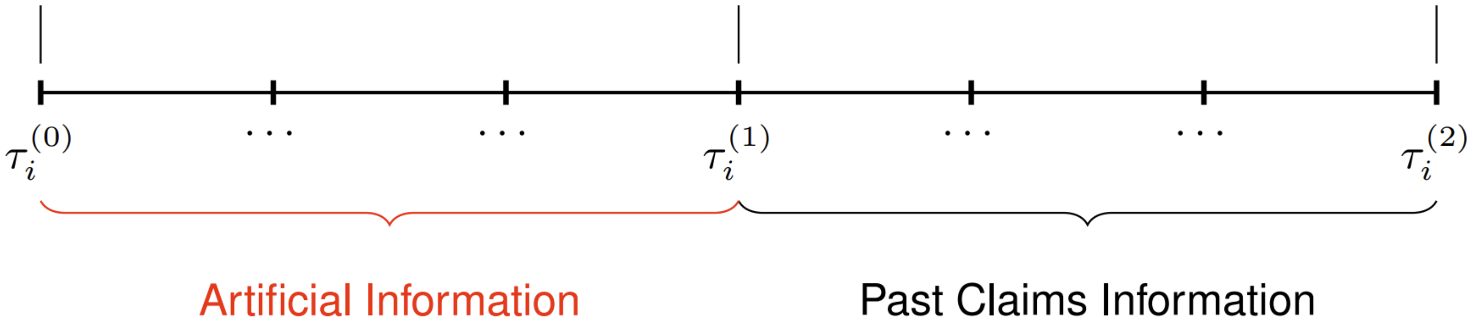
Timeline

The figure bellow illustrates the situation, where the timeline is divided in two sections:

i) The Past Claims Information section: the time period where past claims information is available to compute $n_{i,\bullet}$ and $\kappa_{i,\bullet}$, or the Bonus-Malus level.

ii) The Artificial Information section, from the date of the first insurance contract of insured i) to $\tau_i^{(1)}$.

When the available past claims information is short, experience rating models might be difficult to estimate because the amount of information needed to compute the bonus-malus level, for example, is too small.



Artificial claim history

In the actuarial literature, two methods have been proposed to generate an artificial past claims history:

1- Average outcome

The first method is to suppose that all unobserved years of experience of insured i involve an average expected number of claims $\tilde{\mu}_i$.

This implies a corrected version of $n_{i,\bullet}$ and $\kappa_{i,\bullet}$ for each insured, and consequently a new value of $\ell_{i,t}$ for all t.

2- Most probable outcome

The second method is to suppose that each unobserved year of experience would be considered a year without claims, simply because is the most probable outcome.

This assumption simplifies greatly the computation and the estimation of the BMS because it means using the first method with $\tilde{\mu}_i = 0$ for all insureds.

Reference

This part of the presentation is based on:

J.-P. Boucher (2022). Multiple Bonus-Malus Scale Models for Insureds of Different Sizes. *Risks*, 10(8), 152.

risks

Article

Multiple Bonus–Malus Scale Models for Insureds of Different Sizes

Jean-Philippe Boucher

Departement de Mathematiques, Université du Québec à Montréal, Montréal, QC J3H 5T6, Canada; boucher.jean-philippe@uqam.ca

Abstract: How to consider the *a priori* risks in experience-rating models has been questioned in the actuarial community for a long time. Classic past-claim-rating models, such as the Buhlmann–Straub credibility model, normalize the past experience of each insured before applying claim penalties. On the other hand, classic Bonus–Malus Scales (BMS) models generate the same surcharges and the same discounts for all insureds because the transition rules within the class system do not depend on the *a priori* risk. Despite the quality of prediction of the BMS models, this experience-rating model could appear unfair to many insureds and regulators because it does not recognize the initial risk of the insured. In this paper, we propose the creation of different BMSs for each type of insured using recursive partitioning methods. We apply this approach to real data for the farm insurance product of a major Canadian insurance company with widely varying sizes of insureds. Because the *a priori* risk can change over time, a study of the possible transitions between different BMS models is also performed.

Keywords: claim count; ratemaking; bonus–malus systems; recursive partitioning

Size of each Farm

Remember the predictive expected premium for the Poisson-gamma model:

$$E[N_{i,T}|n_{i,1},\ldots,n_{i,T-1},\boldsymbol{X}_{i,T}] = \lambda_{i,T}\frac{\alpha+\sum_{t=1}^{T-1} n_{i,t}}{\alpha+\sum_{t=1}^{T-1}\lambda_{i,t}}$$

or the weights in the Buhlmann-Straub model, where the random variables were normalised by the *a priori* risk: $Y_{i,t} = \frac{N_{i,t}}{W_{i,t}}$.

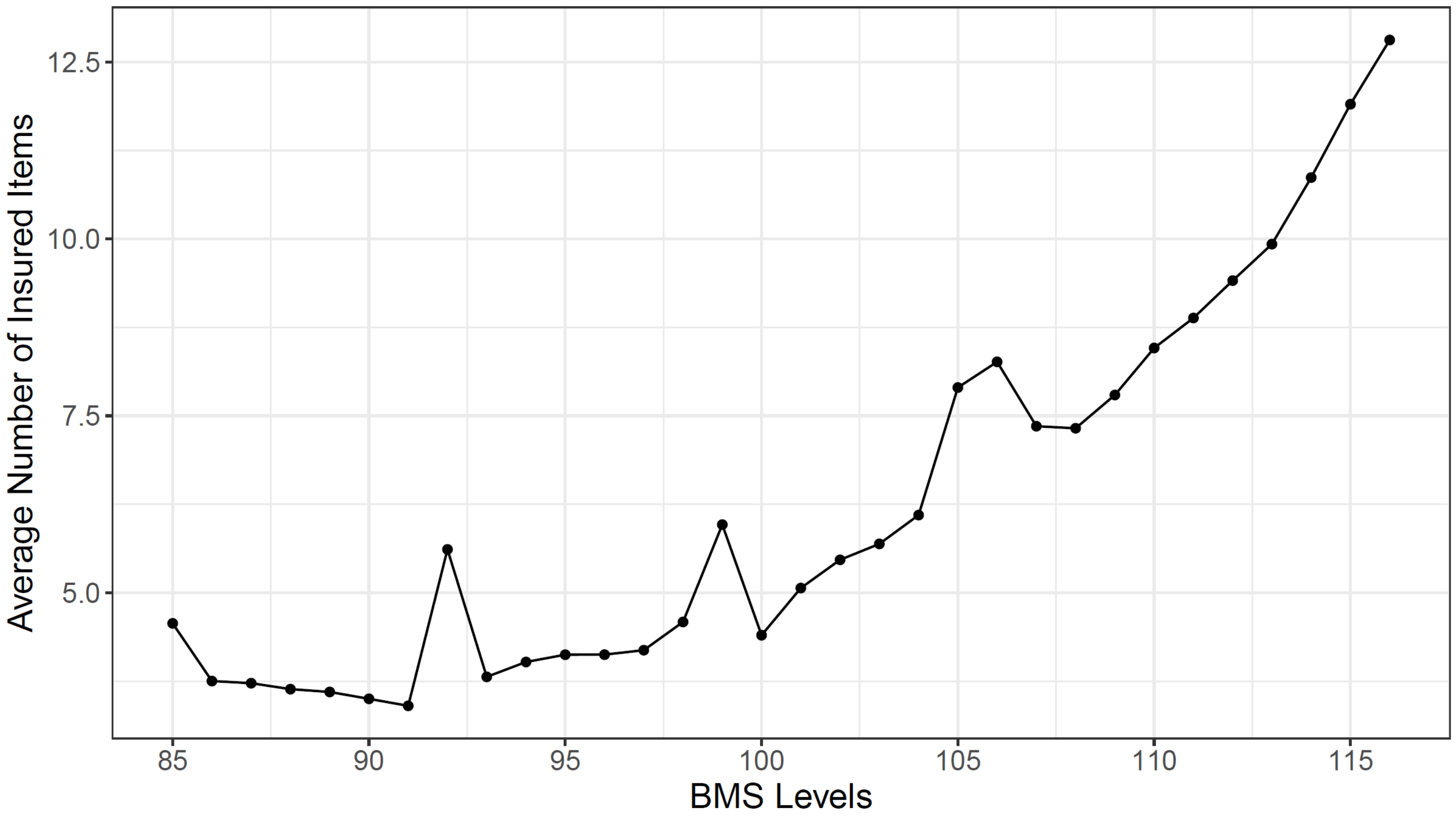
That means that the experience of an insured is *normalized* when it is used in predictive ratemaking. BMS models does not do that: the penalty for a claim does not depend on the *a priori* risk.

Farm insurance

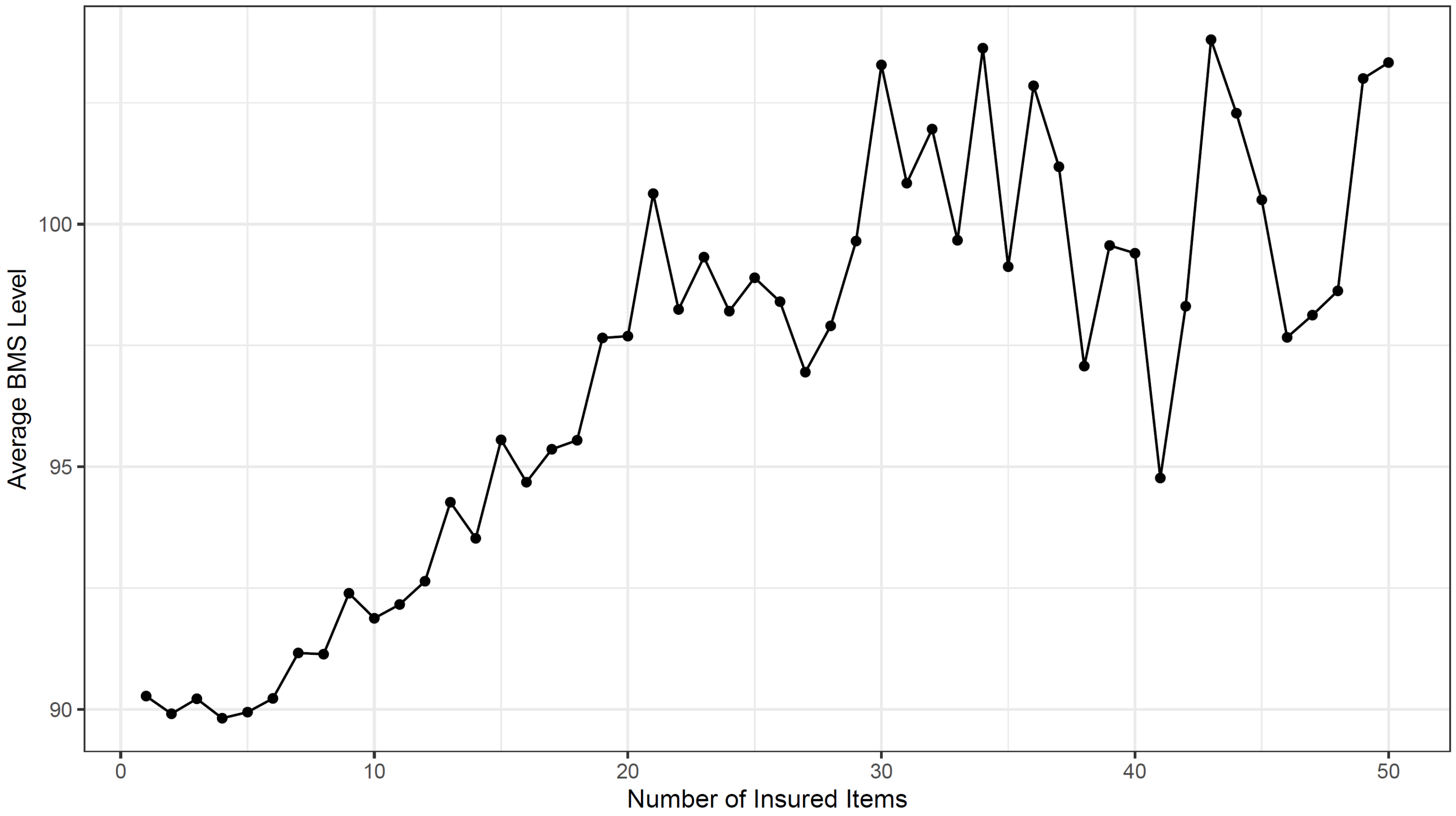
This caused a problem in farm insurance where large farms can be penalized twice:

- In their *a priori* risk;
- With the BMS structure (because they claim more).

Bonus-Malus vs. size of the farm



Bonus-Malus vs. size of the farm



Partitioning the portfolio

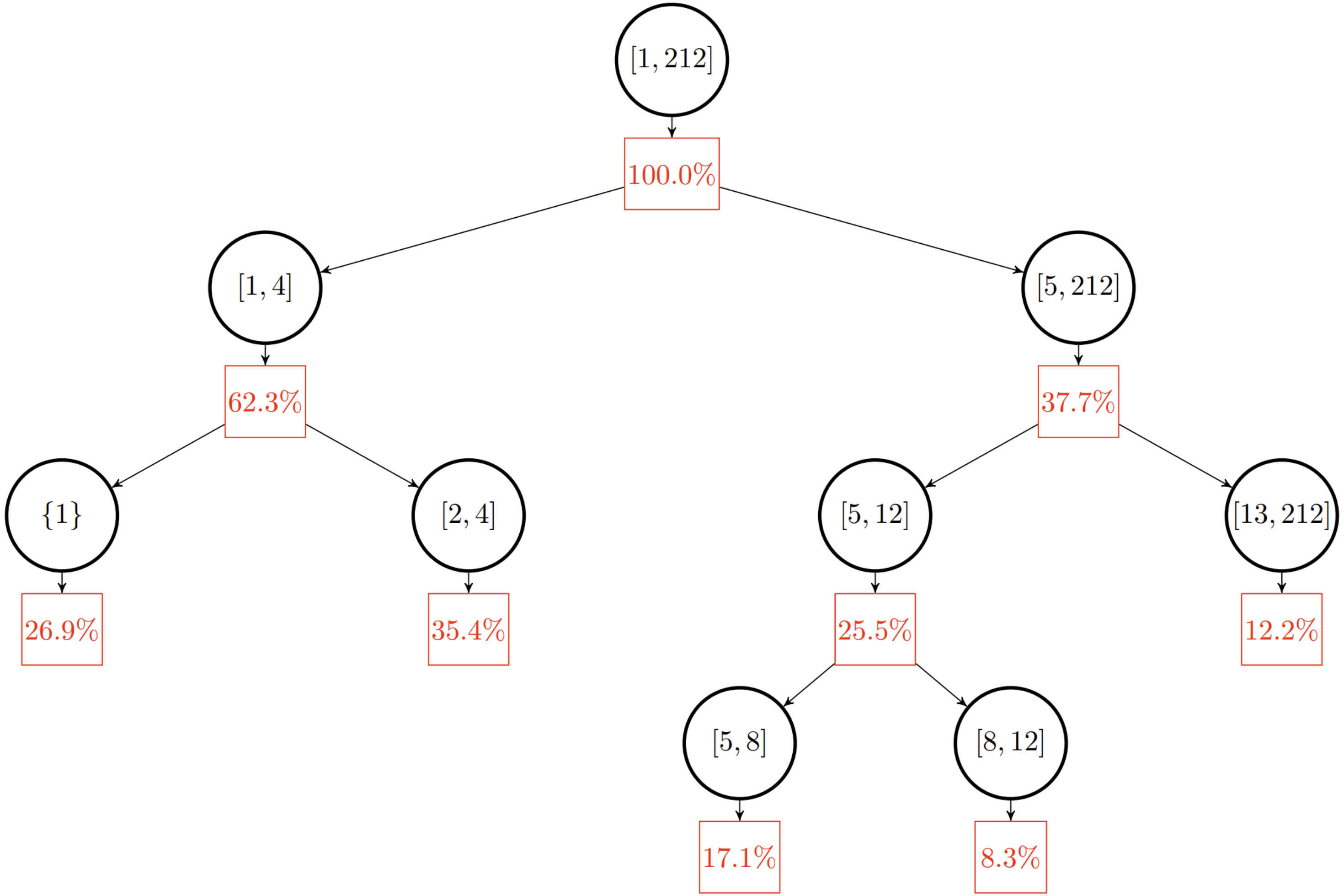
A proposed solution to deal with farms of different sizes was to divide the portfolio into groups. Groups of farms of similar sizes could be created, and each group would have their own experience-rating model, with its own *a priori* rating parameters and its own structural BMS parameters.

Farms could then be more equitably rated, and more correctly rewarded and penalized, as their size would be directly taken into account when performing past claims rating.

Recursive algorithm

To find the best way to group similar farms, a recursive algorithm was proposed.

Partitioning the portfolio by the size



Conclusion

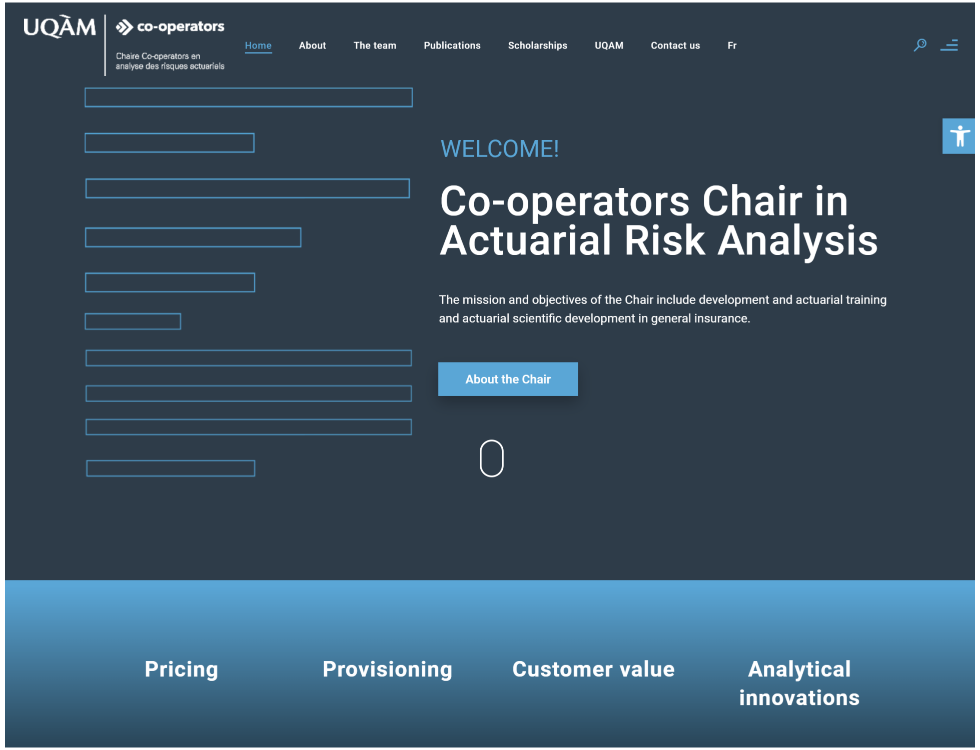
- Bonus-Malus Scales models are already in use in many countries;
- Because of how insurance datasets are now constructed, the current way to calibrate BMS has to be changed;
- We show that iterative GLM approaches can be used to estimate BMS models.
- BMS are really flexible and allows penalty structures that cannot be supposed easily by other models and distributions;
- Despite its simplicity, it has been shown to out-perform many more *advanced* models;

Website of the research Chair

You can check the website of the *Co-operators Chair in Actuarial Risk Analysis* (CARA) for:

- recent publications,
- research projects;
- MSc and PhD fundings,
- etc.:

https://chairecara.uqam.ca/en/

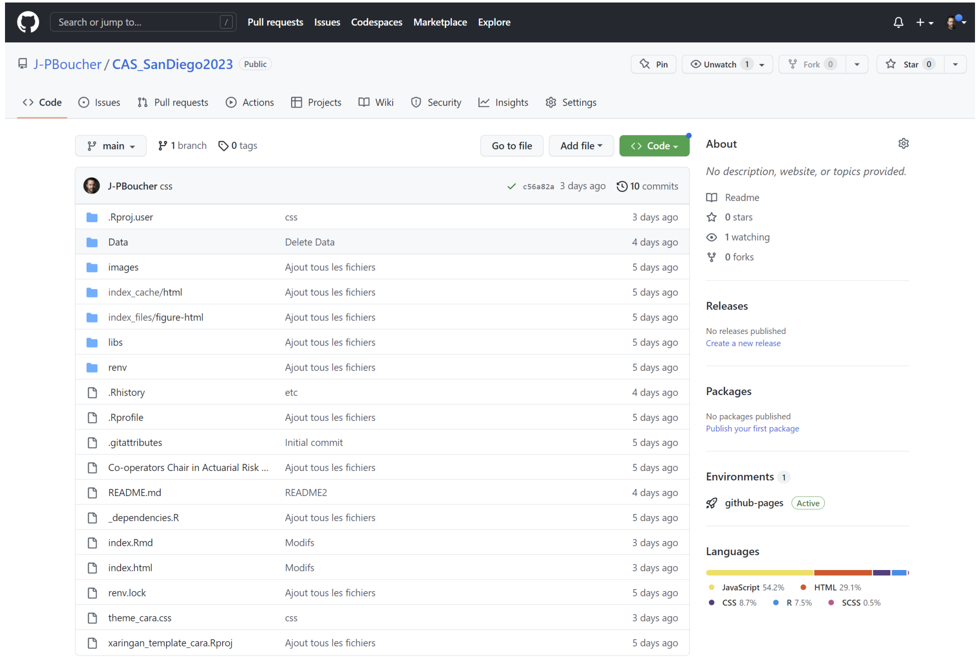


Github

On my github page, you can find:

- this presentation;
- all R script codes used;
- the dataframe *df2.Rda*.

https://github.com/J-PBoucher



Other

References

- J.-P. Boucher (2022). Bonus-Malus Scale Models: Creating Artificial Past Claims History. *Annals of Actuarial Science*, 1-27.
- J.-P. Boucher (2022). Multiple Bonus-Malus Scale Models for Insureds of Different Sizes. *Risks*, 10(8), 152.
- J.-P. Boucher & M. Pigeon (2019), A Claim Score for Dynamic Claim Counts Modelling, *Canadian Institute of Actuaries*.
- J.-P. Boucher & R. Inoussa (2014). A posteriori ratemaking with panel data. *ASTIN Bulletin: The Journal of the IAA*, 44(3), 587-612.

Thanks

Finally, a special thank to my Ph.D. student **Francis Duval** who created this nice *xaringan* template with RMarkdown.