

---

# MULTIPLE BONUS-MALUS SCALE MODELS FOR INSUREDS OF DIFFERENT SIZES

---

A PREPRINT

**Jean-Philippe Boucher**

Chaire Co-operators en analyse des risques actuariels  
Departement de mathematiques  
Universite du Quebec a Montreal  
boucher.jean-philippe@uqam.ca

July 25, 2022

## 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

## 1 Introduction

An experience ratemaking model uses past claims experience to update the basic premium computed with the risk characteristics of the insureds. To avoid penalizing the risky insureds twice, insurers must find a way to consider the *a priori* risks in experience-rating models. Indeed, because risky policyholders already have higher basic premiums than less risky policyholders, they are expected to also claim more often than less risky policyholders. Consequently, they should also be expected to be penalized less for a claim than less risky policyholders. Similarly, risky insureds should be rewarded much more than lower risk insureds if they do not claim.

This problem of treating insureds differently in experience-rating models has long been known in the actuarial community. Indeed, even classic past claim-rating models normalize the past experience of each insured  $i$  before applying claim penalties. For example, by designating  $\lambda_{i,t}$  a measure of the *a priori* risk, for contract  $t$  of insured  $i$ , a normalized past experience  $\frac{\sum_t n_{i,t}}{\sum_t \lambda_{i,t}}$  is used in the Buhlmann-Straub credibility model instead of  $n_{i,\bullet} = \sum_t n_{i,t}$ , which is used in the Buhlmann (1967) credibility model.

Instead on focusing on credibility premiums, other approaches developed joint distributions of all contracts of the same insured to find predictive premiums for claims frequency. For example:

- Models allowing a dynamic time dependence between contracts of the same insured: models based on series of correlated random effects (Bolancé et al. (2007), Abdallah et al. (2016)), jitter models (Shi and Valdez (2014)), time series for count data (Gourieroux and Jasiak (2004), Bermúdez et al. (2018), Bermúdez and Karlis (2021) or Pinquet (2020)),
- Longitudinal models allowing for dependence structures between many type of claims (Abdallah et al. (2016), Pechon et al. (2019a), Pechon et al. (2019b), Gómez-Déniz and Calderín-Ojeda (2018)),
- Longitudinal models allowing dependance between claim frequency and claim severity (Shi et al. (2016), Oh et al. (2020), Jeong and Valdez (2020)).

On the other hand, a Bonus-Malus Scales (BMS) model is an experience-rating model that has been used in practice for a long time (see Lemaire (2012) or Denuit et al. (2007) for classic textbooks, or Frangos and Vrontos (2001), Mehmet and Saykan (2005) or Ágoston and Gyetvai (2020) for research papers on BMS scales). The BMS model is a class-system with a finite number of levels, where a relativity is assigned to each level. Depending on the transition rules of the BMS, insureds usually move down by a level if they do not claim during their contract, and move up a specific number of levels for each claim they make. The insured's new level at the end of the year is then used to compute the next annual premium. To account for the differences in *a priori* risk in automobile insurance, Denuit et al. (2007) uses one BMS model for young drivers and another BMS model for older drivers. However, this approach does not take advantage of the data structure that is now available from insurers. Indeed, to find the relativity of each BMS level, Denuit et al. (2007) uses the classic BMS theory, which is based on aggregate data.

Recently, Boucher and Inoussa (2014) explained how classic BMS theory could be generalized for granular data, where each insured is observed for many contracts. Verschuren (2021) generalizes the BMS approach for multi-products and adds more flexible estimation methods by using the generalized additive model (GAM) theory. Finally, Boucher (2022) shows how BMS models are linked with simple GLM models that have covariates associated with the past claims experience.

Those recent 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, based on the new BMS theory, we show how insurers can create different BMSs to account for the differences in *a priori* risk.

The paper has the following structure. In Section 2, we will carefully explain the recent BMS theory that uses granular data. As an illustration, the BMS model is applied to a farm insurance product. The results are discussed, and the problem related to the *a priori* risk is highlighted. We show that when a single BMS is used on the entire insurance portfolio, higher BMS levels are filled with bigger farms, and similarly, we show that bigger farms have a higher average BMS level than smaller farms. In Section 3, we show how to group farms of similar size to create distinct and independent BMS models. This more equitably rates the farms, and better rewards and penalizes them, because their size is directly taken into account for past claims rating. To find the best groups to form among the farms, similar to what is done with classification trees, we use a recursive partitioning algorithm (see Diao and Weng (2019) for a recent use of classification trees in a credibility theory situation). Because the size of the farm, which is defined by the number of insured items, can change at any time, a study of the transitions between different BMS models is also performed. The last section concludes.

## 2 Review of the BMS Model

## 2.1 Summary of the Past-Claims-Rating Model

Boucher (2022) defines two kinds of variables to use in experience rating:

1. The variable to model, named the **target** variable;
2. The information used to define what we consider the past claim experience, named the **scope** variable.

For this paper, we will model the number of claims (i.e. the target variable), based only on the number of past claims (i.e. the scope variable). In other words, the idea of the experience rating model of this paper is to predict the number of claims of insured  $i$ , for contract of period  $T$ , by taking into account all the past number of claims, from contract  $t = 1$  to contract  $T - 1$ . A new insured with  $T = 1$  is someone without any past insurance experience. Formally, we are looking to compute the conditional expected value of  $N_{i,T}$ , which is the number of claims of insured  $i$  for contract of period  $T$ , defined as :

$$E[N_{i,T} | \mathbf{n}_{i,(1:T-1)}, \mathbf{X}_{i,T}] \quad (1)$$

where  $\mathbf{n}_{i,(1:T-1)}$  is a vector of all number of claims of each contract between time 1 and time  $T - 1$  for insured  $i$ , and  $\mathbf{X}_{i,T}$  is a vector containing covariates used in the ratemaking, for contract  $T$ . This usually corresponds to information about the age of the insured, the marital status of the insured, etc.

## 2.2 The Bonus-Malus Scale Models

To model the number of claims of insured  $i$ , for contract  $T$ , the Poisson distribution of mean  $\lambda_{i,T}$  is usually the starting point, but other count distributions can easily be used (see Denuit et al. (2007) for a review of using count distributions to model the number of claims, or Cameron and Trivedi (2013) for an exhaustive review of count distributions). For this paper, for the sake of simplify, we will only use the Poisson distribution.

For experience rating, to differentiate between new insureds and insureds with experience, past-claims-rating models use both the number of past claims and the number of contracts without claims, defined as  $\kappa_{i,t} = I(n_{i,t} = 0)$ . In such situations, the mean parameter of the Poisson distribution can be expressed as:

$$\lambda_{i,T} = \exp(\mathbf{X}'_{i,T}\boldsymbol{\beta} + \gamma_0(100 - \kappa_{i,\bullet}) + \gamma_1 n_{i,\bullet}) \quad (2)$$

where, for insured  $i$ ,  $n_{i,\bullet} = \sum_{t=1}^{T-1} n_{i,t}$  corresponds to the insured's total number of past claims, and  $\kappa_{i,\bullet} = \sum_{t=1}^{T-1} \kappa_{i,t}$  is the sum of policy periods without a claim. The same mean parameter can be expressed as:

$$\lambda_{i,T} = \exp(\mathbf{X}'_{i,T}\boldsymbol{\beta} + \gamma_0(100 - \kappa_{i,\bullet} + \frac{\gamma_1}{\gamma_0} n_{i,\bullet}) = \exp(\mathbf{X}'_{i,T}\boldsymbol{\beta} + \gamma_0 \ell_{i,T}), \quad (3)$$

with:

$$\ell_{i,T} = 100 - \kappa_{i,\bullet} + \frac{\gamma_1}{\gamma_0} n_{i,\bullet}$$

where  $\gamma_0$  is the **Relativity** parameter, and  $\Psi = \frac{\gamma_1}{\gamma_0}$  is the **Jump** parameter. The new variable  $\ell_{i,T}$ , based on  $\kappa_{i,\bullet}$  and  $n_{i,\bullet}$ , summarizes all past claim experience and is called a **claim score**. This approach is called the *Kappa-N* model (Boucher (2022)), and because it is defined by the mean parameter, it can easily be used with many count distributions. The Kappa-N model can be interpreted as follows:

- For an insured  $i$  without insurance experience, we have  $n_{i,\bullet} = 0$  and  $\kappa_{i,\bullet} = 0$ , which means that a new insured without experience will have a claim score of 100, i.e. an **entry level** of 100.

Insured $i$	Years										$-\kappa_{i,\bullet}$	$n_{i,\bullet}$
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020		
1	0	0	0	0	0	0	0	0	0	0	-10	0
2	2	0	1	0	0	0	2	0	1	0	-6	6
3	4	1	2	0	0	0	0	0	0	0	-7	7

**Table 1:** Insureds with claims experience

- Each year without a claim will decrease the claim score by 1;
- Each claim will increase the claim score by  $\Psi$ .
- The impact of a single claim on the premium is equal to  $\Psi$  years without claims.
- The penalty for a claim is  $(\exp(\Psi\gamma_0) - 1)\%$ .
- Each year without a claim decreases the premium by  $(1 - \exp(-\gamma_0))\%$ .

Boucher (2022) showed that the prediction quality of the Kappa-N model is significantly better than standard count models that do not use any covariates linked to experience rating. However, the Kappa-N model has some major flaws. The model does not limit the claim score to minimum or maximum values. In some situations, it can also lead to extreme surcharges for insureds who claimed frequently. Moreover, because it is based on  $\kappa_{i,\bullet}$  and  $n_{i,\bullet}$ , the Kappa-N model does not allow for any forgiveness and the model does not weight past claims by their age (Bolancé et al. (2007)).

### 2.2.1 Numerical Example

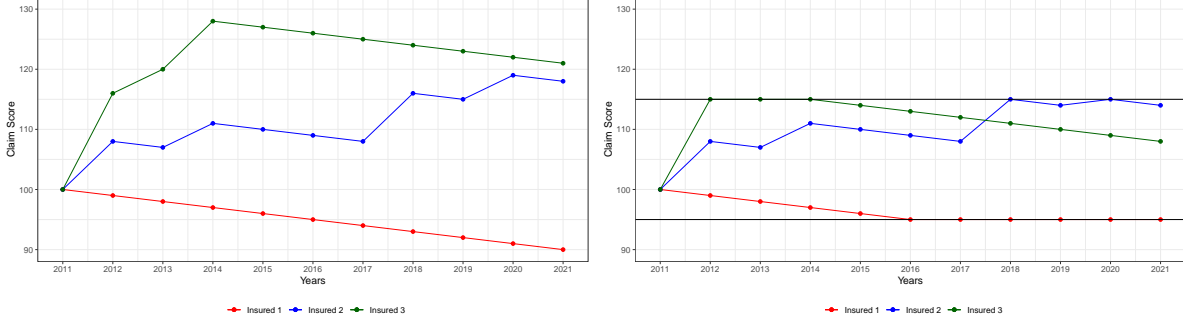
Suppose that we have the insureds from Table 1, where the insurance experience is shown for three insureds. Suppose a Kappa-N model was fitted, resulting in a rating structure that implies a decrease of one level for each year without a claim. The estimated jump-parameter  $\Psi$  is 4, meaning that each claim penalizes the insured with an increase of 4 to the claim score. If we start at level 100 in 2011, because of the values of  $\kappa_{i,\bullet}$  and  $n_{i,\bullet}$  (also shown in Table 1), insureds would respectively be at levels 90, 118 and 121 in 2021. The graph on the left of Figure 1 shows another way to see how the level  $\ell_{i,2021}$ , for insured  $i = 1, 2, 3$ , are computed.

If we want to weight claims by their age and introduce some form of forgiveness to the rating model, a solution is to limit the value of all claim scores, for all past insurance contracts. For illustration, we suppose that 115 is the maximum value of the claim score, i.e  $\ell_{max} = 115$ , and the minimum value is 95, i.e  $\ell_{min} = 95$ . To compute premiums in 2021, we apply these two limits to all past claim scores. Figure 1 shows how the limits applied to past contracts impact the claim score in 2021 for all three insureds. With this approach, we see that many of the problems mentioned earlier are at least partially solved: past claims can ultimately be forgiven, and the model introduces a form of time weighting for past claims.

By adding maximum and minimum claim-score values for past contracts, the Kappa-N approach developed so far has been transformed into what is usually called a **Bonus-Malus Scale** system (BMS), where the claim score can be seen as a **BMS level**, or simply a BMS score. More formally, the BMS is a class system with a finite number of levels (when  $\Psi$  is an integer) where a relativity is assigned to each level. For this BMS model, the transition rule is simple: an insured moves down by one level if there is no claim over the course of the contract, and moves up by  $\Psi$  levels for each claim.

BMS models can be defined using four structural parameters: the entry level  $\ell_0$ , the jump parameter  $\Psi$ , the maximum level of the system  $\ell_{max}$  and the minimum level of the system  $\ell_{min}$ . For a specific insured, the BMS level is defined as:

$$\ell_{i,t+1} = \ell_{i,t} - \kappa_{i,t} + \Psi \times n_{i,t}, \quad \text{with } \ell_{min} \leq \ell_{i,t} \leq \ell_{max},$$



**Figure 1:** Insureds with claim experience, with and without limits  $\ell_{min}, \ell_{max}$

where the level  $\ell_{i,t}$  is always between  $\ell_{min}$  and  $\ell_{max}$  for all  $t = 1, \dots, T$ .

For the Kappa-N model, estimating the parameters by maximum likelihood is simple. However, for BMS models, because the model needs to recompute the claim score path of each insured from their first contract, finding the best values for structural parameters  $\Psi, \ell_{min}$  and  $\ell_{max}$  is not that easy. To estimate the structural parameters, a grid search for all structural parameters can be performed. That means that all values are tried, and we select the BMS model with the biggest log-likelihood. By limiting the possible values of  $\Psi, \ell_{min}$  and  $\ell_{max}$ , this approach is not difficult, nor is it too time consuming.

### 2.3 Summary of the Numerical Illustration

#### 2.3.1 Data Used

We use a sample of the farm insurance data from a major insurance company in Canada, using contracts between 2010 and 2020. The general form of the data used is shown in Table 2, where each line in the database corresponds to a specific coverage in an annual contract. For each observation, we have information about the insured, the contract, the items covered, but we also see the date of the first insurance contract with the insurer. Information about claims that happened during that period of time is also available.

Policy Number	Number of Items	Effective Date	First Insurance	Coverage	...	Province	Number of Claims	Costs of Claims
...	...	...	...	...	...	...	...	...
125721	2	2017-01-15	1995-01-15	MACHINERY	...	Ontario	2	186,592
125722	15	2017-03-22	2013-03-22	MACHINERY	...	Quebec	0	0
125723	1	2016-01-11	1993-11-05	MACHINERY	...	Manitoba	1	18,889
125724	27	2018-02-17	2018-02-17	MACHINERY	...	Nova Scotia	1	7,444
...	...	...	...	...	...	...	...	...

**Table 2:** Fictive Data Sample - Contract Level

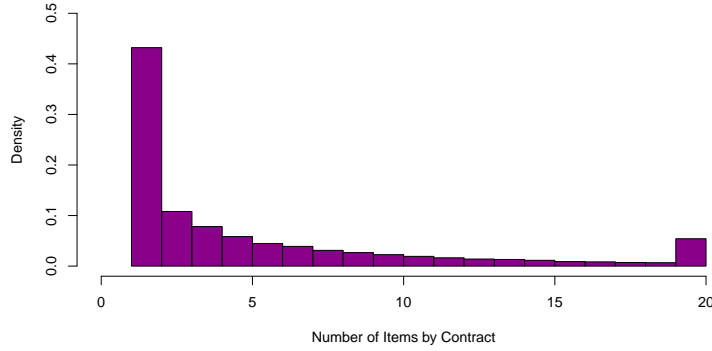
Unlike with automobile insurance in Canada, where insureds frequently move from one insurer to another, farm insurance has relatively stable insureds. In our dataset, the average number of years with the insurer is 18.4, and the maximum observed years is close to 60<sup>1</sup>. The maximum available years of past claims experience for all insureds is 15 years, and only insurance experience with the insurer is available. We considered the first year of insurance of any insured to be the first year with the insurer. In other words, if a farm is first seen in the database in 2003, this farm will

<sup>1</sup>Farms are sometimes passed from generation to generation. Insurance experience would not be reset in such a case.

Distributions	BMS Parameters				Log-likelihood (train)	Log. Score (test)
	$\ell_{max}$	$\ell_{min}$	$\hat{\Psi}$	$\hat{\gamma}_0$		
Poisson	116	85	6	0.0312	-8,490.026	2,857.029

**Table 3:** Results of the Poisson BMS model

be considered a new insured without any prior experience in 2003.



**Figure 2:** Distribution of the number of items by contract

For the farm insurance product, an item corresponds to a specific tractor or combine which specific information is available. With a total of approximately 700,000 insured items insured for more than 120,000 contracts, the average number of items insured per contract is around 6. The distribution of the number of items insured per contract can be seen in Figure 2. Almost 50% of all farms only have one insured items, while approximately 10% of farms have more than 20 insured items. More precisely, 40 farms have more than 100 insured items, with a maximum of more than 200 insured items for a single contract. As we will see in the next sections, the difference between small farms and larger farms is important for BMS models.

### 2.3.2 Estimated Parameters of the BMS Model

The estimated parameters of the Poisson BMS model are shown in Table 3. The log-likelihood value is shown. For the test dataset, the logarithmic score defined as  $\sum_{i=1}^n -\log(\Pr(n_i; \hat{\lambda}_i))$  has been used (see Roel et al. (2017) for details or descriptions of other scores) to define the prediction quality. To better understand the results obtained for the Poisson BMS model, we can compute the discounts and surcharges of the model, based on the number of past claims. More concretely, we then have:

- The jump parameter  $\Psi$  is equal to 6, meaning that each claim increases the BMS level by 6. After a claim, an insured would need 6 years without a claim to return to the original premium (before the claim).
- The value of  $\gamma_0$  is 0.0312. That means that the penalty for a claim is equal to  $\exp(0.0312 \times 6) - 1 = 20.6\%$ , and each year without claim decreases the premium by  $1 - \exp(-0.0312) = 3.07\%$ .
- The maximum BMS level is  $\ell_{max} = 116$ , meaning that the maximum surcharge, compared to level 100, is  $\exp(0.0312 \times 16) - 1 = 64.7\%$ ;
- The minimum BMS level is  $\ell_{min} = 85$ , meaning that the minimum surcharge, compared to level 100, is  $1 - \exp(-0.0312 \times 15) = 37.3\%$ .

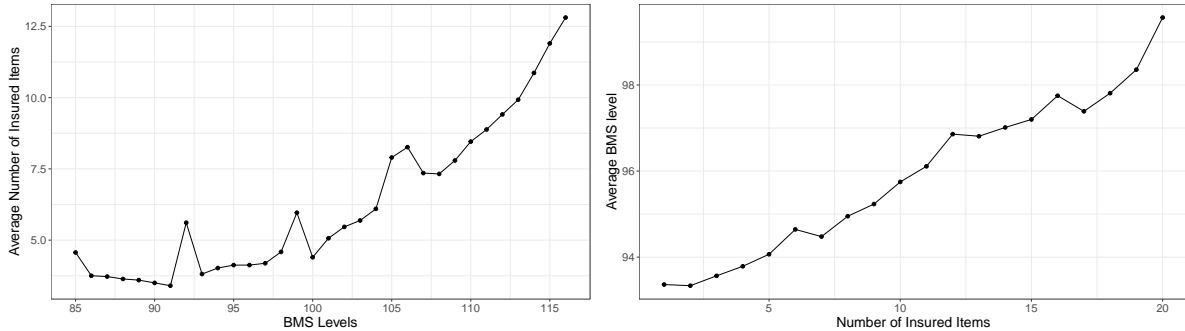
As we can see, these basic results are found and computed easily. This method of computing the surcharges and discounts would clearly be useful to any insureds, brokers or administrators. It is simple to explain to insureds how large their penalties for a claim will be, and how long they will be penalized for that claim. Another interesting result of the BMS model is that all insureds will have a premium located between 0.627 and 1.647 times the basic premium for a new insured at level 100. This narrowly limits the range of premiums.

### 2.3.3 Problems With the Size of Farms

By comparing the predicted and the observed claims frequency on the train and the test datasets, Boucher (2022) showed that the BMS model seems to fit the data well. We see that classifying insureds by their claim score (or BMS level) works well as the insureds with higher levels has worse claims experience than insureds with lower levels.

However, the size of each farm in the insurance portfolio is different and size has a direct impact of the past rating model. Figure 2 showed the distribution of the number of insured machinery (called items) per farm. The BMS model used here generates the same surcharges and the same discounts for all insureds. However, because large farms are expected to have more claim than smaller ones, they should normally also be expected to be penalized less for a claim. Similarly, a large farm should be rewarded much more for a year without claim. An experience-rating system that does not recognize this type of situations may appear to penalize larger farms twice.

That means that the connection between BMS levels and the size of the farm is noteworthy. To more clearly see the impact of the number of insured items on each farm, using the BMS model from Table 3, we compute the BMS level of each contract in the database. For each BMS level, we compute the average number of insured items on each farm for that level. Similarly, we computed the average BMS level based on the number of insured items. Figure 3 illustrates the result. As we expect, the higher BMS levels are filled with bigger farms, and bigger farms will have an average BMS level much higher than smaller farms.

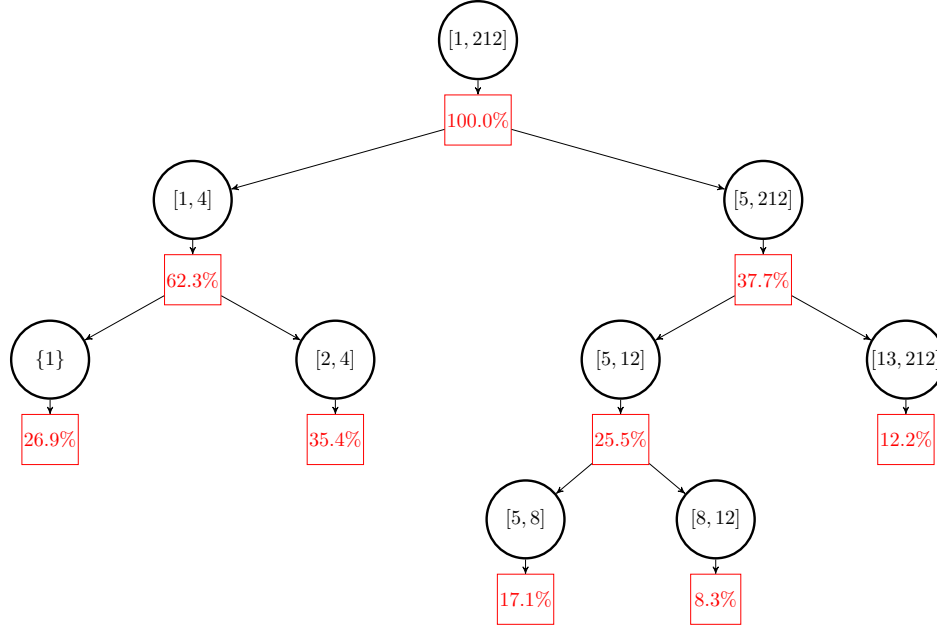


**Figure 3:** Average number of insured items by BMS level (left) and average BMS level by number of insured items

Despite the prediction quality of the BMS model, the BMS model could seem unfair to many insureds and regulators because it does not recognize the initial risk of the insured. To correct this situation and promote the use of the BMS in practice, the BMS model should be generalized to include the *a priori* risks in the rating structure.

## 3 Partitioning the Portfolio

As proposed by Denuit et al. (2007) for young drivers in Belgium, a way to deal with farms of different sizes would be to divide the portfolio into groups. Indeed, 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.



**Figure 4:** Recursive Partitioning by the Number of Insured Items

To find the best groups to form among the farms, similar to what is done with classification trees, we perform a recursive partitioning of the portfolio (see Diao and Weng (2019) for more details, but also for an application to compute more accurate credibility premiums) Because we expect the risk to be proportional to the number of insured items in each contract, we need to find a division point based on the number of items from which we create two groups: insureds with a smaller (or equal) number of items than the point division, and another group with more items than the division point.

We use an algorithm to test each possible sub-group of our insurance portfolio by cross-validation. We split the training dataset into five folds, from which four of the five are used to estimate the parameters of a Poisson BMS model, and the remaining fold is used to compute predictive statistics. To evaluate the prediction quality on the test fold, because we are working with count data, the logarithmic score defined earlier has been used. For each item value from 1 to 212 (the maximum number of items observed in a single contract), we are looking for the value where the sum of the logarithmic score of the two groups is maximized. Then, for each of the two sub-groups created, we create two more new groups that generate a better total logarithmic score. We repeat this until the sum of all scores cannot be improved, or when a group is composed only of farms with the same number of insured items.

The resulting classification tree of the number of items is shown in Figure 4, where the percentage shown corresponds to the proportion of contracts that belong to that interval of the training dataset. We can see from the algorithm that:

1. The first iteration divides the insureds into two groups: insured with four or less than four insured items, which represents 62.3% of all insureds; and insureds with more than four items, which represents 37.7% of the insureds.
2. The first group of insureds is then divided again into two groups: farms with one insured item (26.9%), and farms with between two and four insured items (35.4%). The second group is also divided into two more groups: between five and 12 items (31.0%) and insureds with more than 12 items (6.7%).
3. Finally, only the group with between five and 12 items can be divided again: groups with between five and eight, and between nine and 12 items are created.



### 3.1 Analyzing Each Group of Farms

To better understand the division of the portfolio into groups, Table 4 shows a descriptive summary of each group of farms from division step 0,1,2 and 3. For each division, we observe major differences between each group of farms. For example, the average claims frequency of the portfolio is 2.36%, but it is 0.36% for insureds with only one insured item and is slightly under 10% for farms with more than 13 insured items. Past claims experience, summarized by variables  $\bar{\kappa}_\bullet$  and  $\bar{n}_\bullet$ , also shows the difference between groups: on average, small farms have fewer reported claims (0.49) than the larger farms (1.68). The number of years without a claim, expressed by  $\bar{\kappa}_\bullet$ , is much closer for each group, except for the smaller farms with only one insured item. It is also interesting to see that larger farms have more insurance experience, noted  $\bar{\tau}$ , than small ones. Generally, the larger the farm is, the longer it stays with the insurer.

Depth	Division Group	Number of Items	Proportion of Contracts	Average Nb. of Items	Claims Frequency	Past Claims History		
						$\bar{n}_\bullet$	$\bar{\kappa}_\bullet$	$\bar{\tau}$
0	0A	[1, 212]	100.0%	5.85	2.36%	0.83	10.30	11.08
1	1A	[1, 4]	62.3%	2.00	0.71%	0.58	9.95	10.50
	1B	[5, 212]	37.7%	12.22	5.08%	1.24	10.88	12.03
2	2A	{1}	26.9%	1.00	0.36%	0.49	9.53	10.00
	2B	[2, 4]	35.4%	2.75	0.98%	0.64	10.26	10.87
	2C	[5, 12]	25.5%	7.57	3.25%	1.04	10.99	11.96
	2D	[13, 212]	12.2%	21.94	8.91%	1.68	10.64	12.17
3	3A	{1}	26.9%	1.00	0.36%	0.49	9.53	10.00
	3B	[2, 4]	35.4%	2.75	0.98%	0.64	10.26	10.87
	3C	[5, 8]	17.1%	6.24	2.63%	0.94	10.94	11.83
	3D	[9, 12]	8.4%	10.29	4.52%	1.23	11.09	12.24
	3E	[13, 212]	12.2%	21.94	8.91%	1.68	10.64	12.17

*Table 4: Descriptive summary of each group of farms*

Covariates based on the characteristics of the farm have been used in the claim count model. More specifically, by referring to equation (2),  $\mathbf{X}_{i,T}$  is a  $7 \times 1$  vector, with six covariates and an intercept. For privacy reasons, but also because this is not the focus on the paper, covariates are simply labelled as  $X_1 - X_6$  and their meaning is not explained. Figure 5 shows the mean of each covariate used in all count models for all five groups, with the dashed line representing the average for the whole portfolio. Major differences between groups can be observed for covariates  $X_1$  and  $X_2$ . This is not surprising because those covariates are directly linked to the number of insured items per farm. Covariates  $X_3$ ,  $X_4$ ,  $X_5$ , and  $X_6$ , on the other hand, do not vary depending on the group.

Figure 6 shows the distribution of each group of farms of division 3 across all BMS levels, for the original BMS models summarized in Table 3. The conclusion of that figure is similar to what we saw with Figure 3: the proportion of larger farms in the highest levels of the BMS is greater than the proportion seen in the lower BMS levels. We see, for example, that the proportion of contracts from groups 3A and 3B is smaller in level 115 than in level 85. The opposite is observed for group 3E. However, the proportion of insureds from groups 3C and 3D seems to be stable across all BMS levels.

### 3.2 Different BMS Models versus the Original BMS Model

By dividing the portfolio into groups based on the number of insured items, we also obtained five different BMS models for each group of farms. Results of the estimation of each BMS model are shown in Table 5. A comparison with the original BMS, which has been fitted to the whole portfolio, is also shown, where the log-likelihood statistics and logarithmic scores of the original BMS model is divided by group.

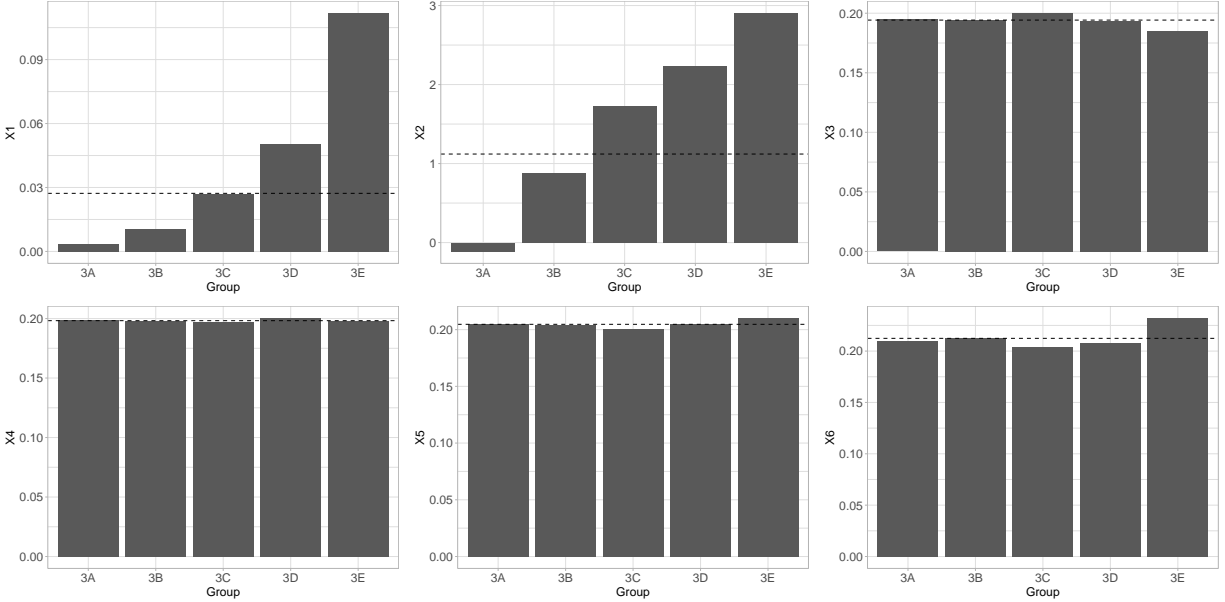


Figure 5: Mean of all covariates used, by groups (dashed line for the mean of the whole portfolio)

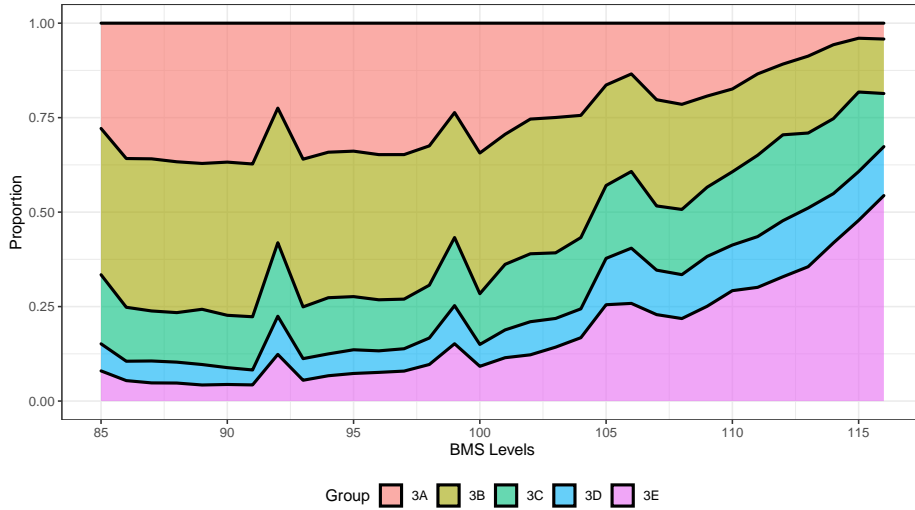


Figure 6: Proportion of each groups within BMS levels for the BMS model

Because it generalizes the original BMS model, a specific BMS model for each group obviously improves the log-likelihood of each group and the total log-likelihood ( $-8490.03$  vs.  $-8410.47$ , almost 80 points, for 33 additional parameters). Improving the log-likelihood that much could be a sign of overfitting. However, this is not the case because, interestingly, the total logarithmic score used on the test dataset also improved ( $2857.03$  vs  $2843.08$ ), and we can see that the logarithmic score of each group improved, except for group 3B. This is an indication that an experience rating model that uses the size of the insured should be considered in farm insurance.

We can also see that the estimators of the structural parameters  $l_{max}$ ,  $l_{min}$  and  $\Psi$  are very different from one group to another. This can be interpreted in many ways. Using the estimated value of  $\Psi$ , we can see how many years it takes to forgive one claim for each group. Another place we can see the impact of the BMS parameters is on the spectrum of possible premiums, which is very different for each group. For example, the maximum/minimum BMS

Group	Original BMS						BMS by group					
	$\hat{\ell}_{max}$	$\hat{\ell}_{min}$	$\hat{\Psi}$	$\hat{\gamma}_0$	Loglik.	Log.Score	$\hat{\ell}_{max}$	$\hat{\ell}_{min}$	$\hat{\Psi}$	$\hat{\gamma}_0$	Loglik.	Log.Score
3A	116	85	6	0.0312	-554.37	114.57	156	99	7	0.0713	-539.06	107.70
3B	116	85	6	0.0312	-1,655.34	550.50	114	85	11	0.0425	-1,633.46	553.19
3C	116	85	6	0.0312	-1,778.24	563.72	118	85	6	0.0320	-1,768.57	558.16
3D	116	85	6	0.0312	-1,336.60	459.63	120	88	6	0.0311	-1,320.47	457.57
3E	116	85	6	0.0312	-3,165.47	1,168.61	107	85	4	0.0431	-3,148.91	1,166.47
Total	116	85	6	0.0312	-8,490.03	2,857.03	.	.	.		-8,410.47	2,843.08

**Table 5:** Comparison of the original BMS and the BMS by group model

levels for group 3A are 156/99 (57 levels), while the maximum/minimum BMS levels for group 3E are 107/85 (22 levels).

Group	Discount by year without claim	Surcharge by claim	Claim Impact*	Maximum Surcharge	Maximum Discount
0A	3.08%	20.6%	24.5%	65.0%	37.5%
3A	6.88%	64.7%	76.9%	5321%	6.90%
3B	4.16%	59.6%	66.5%	81.3%	47.1%
3C	3.15%	17.4%	21.2%	78.0%	38.2%
3D	3.06%	20.5%	24.3%	92.0%	31.1%
3E	4.22%	18.8%	24.1%	35.2%	47.6%

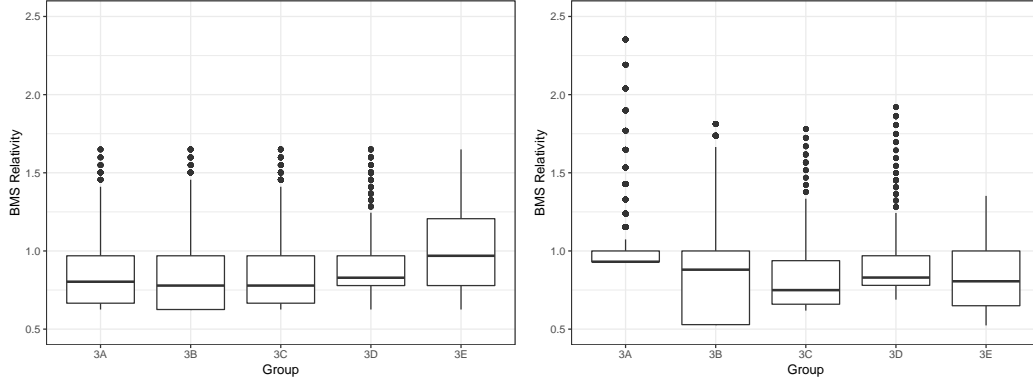
**Table 6:** Discounts and surcharges for each group

Using the structural parameters of the BMS also allows us to compute various discounts and surcharges, as shown in Table 6. The table shows the impact of a year without a claim and the surcharge for each claim. There are only two possibilities for the insured: to claim or not to claim; thus we have to compare relativities between a scenario where the insured claims, and a scenario where the insured does not claim. This is shown in the Claim Impact column. Finally, based on the estimated values of  $\ell_{max}$  and  $\ell_{min}$ , the maximum surcharge and the maximum discount for each BMS model is also shown. Several observations can be made about these tables:

- Group 3A, which is composed of farms with only one insured item, seems very different than the other farms. Perhaps these small insureds are family farms, while the farms in the other groups are more industrial.
- Group 3A generates extreme surcharge values for insureds who claim too much: they are more than 53 times the basic premium for insureds at the top of the bonus-malus scale. Even if this BMS model seems to be better than the original BMS model for group 3A (when looking at the logarithmic score), this is not realistic and cannot be applied in practice. On the other hand, groups 3B, 3C and 3D have similar maximum surcharges ranging from 0.780 to 0.920, which can also be seen as an indication that small farms with one insured item seems very different than the other farms. The maximum surcharge for larger farms is limited to an increase of 35.2% compared to the basic level (100).
- The maximum discounts are similar for each group except group 3A. Compared to a new insured at level 100, insureds from group 3A can expect a maximum discount of only around 7%, but other insureds can obtain discounts ranging from 31.1% to 47.1%.
- The surcharge for each claim is very different from one group to another. We see that for group 3A and 3B, the impact of a claim is similar, while the impacts are similar for groups 3C, 3D and 3E.

To more clearly understand the difference between having five different BMS model, one for each group, and a single BMS model that is applied directly to the whole portfolio, the left graph of Figure 7 shows the distribution of the BMS

relativities for all groups with the single BMS model when it is applied to the whole portfolio. As seen in Table 6, the maximum relativity of the original BMS model is 1.65, and the minimum is 0.625 (1 – 37.5%), but the figure also shows how the contracts are spread across all levels, for each group. The graph on the right of the same figure shows the same result but for five different BMS models. Group 3A includes extreme cases where the BMS relativities are much higher than 2.5.



**Figure 7:** Distribution of the BMS relativities for each group (left: original BMS, right: BMS by group)

Finally, an analysis was also performed of the estimated *a priori* parameters related the covariates used in the models. Table 7 shows some similarities between the values of the parameters, but also large differences between groups.

	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	$\hat{\beta}_6$
0A	0.931	-0.173	-0.176	0.091	0.052	0.008
3A	0.976	-0.953	-0.719	0.141	0.157	-0.133
3B	0.888	-0.799	0.170	0.009	0.247	0.350
3C	0.892	-0.864	-0.268	0.026	0.014	-0.031
3D	0.828	-1.099	-0.045	-0.010	0.159	-0.012
3E	0.877	-0.445	-0.252	0.167	-0.022	-0.045

**Table 7:** Estimated parameters for BMS models applied to each group

The results obtained by dividing the farm insurance portfolio into subgroups to take into consideration the size of the farms is an interesting way to deal with the size of the insureds when conducting experience rating. Indeed, even if the final model is much more complex, and has many structural parameters that results in a cumbersome approach to estimating all parameters, the final results improve prediction capacity compared to the standard BMS model (a.k.a. the original BMS model). Dividing farms by the number of insured items is an interesting way to evaluate the impact of the size of a farm. Other variables or covariates can be used, instead of the number of insured items, to evaluate the insured’s *a priori* risk. In automobile insurance, the driver’s age could be used to divide the insureds into subgroups. For other products, dividing insureds into groups might work if we can find heterogeneous groups of insureds in the portfolio.

### 3.3 Limits of the Approach

Longitudinal count models, such as those seen in Abdallah et al. (2016) or Pechon et al. (2019a), can directly use the number of farms as the exposition basis. Consequently, those models allow for the computation predictive premiums that consider the *a priori* risk while using much less parameters than the classification tree BMS model proposed here. However, as mentioned by Boucher (2022), these longitudinal models are often difficult for insurers to implement and

BMS solutions is often be required.

The classification tree BMS model divides the portfolio and independently estimates all the BMS parameters of each group. That means that we consider each group to be fundamentally different, which is not the case. They all have something in common: small farms might be a little bit different than bigger farms, but ultimately, the same types of machinery are covered and they face the same kinds of perils (thefts, hail, fire, etc.). A way to correct this flaw could be to allow the sharing of some  $\beta$  parameters between each group. This solution will reduce the number of parameters of the model, which often leads to a reduction of overfitting. This could be an idea for future researches because it needs developing another estimating algorithm.

### 3.3.1 A Study of Transition Rules

Dividing the portfolio into different groups causes a bigger problem. Indeed, nothing is preventing a farm from buying or selling items (tractors) at any moment in time. That means that, for example, a farm in group 3B could transition to group 3A simply by removing a tractor in its insurance contract. Because we do not have access to past risk characteristics (in other words the number of items insured in the past), the past claims experience of that farm will then be used with a BMS structure with  $\{\ell_{max}, \ell_{min}, \Psi\} = \{156, 99, 7\}$ , instead of  $\{114, 85, 11\}$ . Transitioning to other groups has obviously not been considered in the modeling, and it leads to illogical results, where, for example, the insured will receive a new surcharge for past claims even if no new claim was reported.

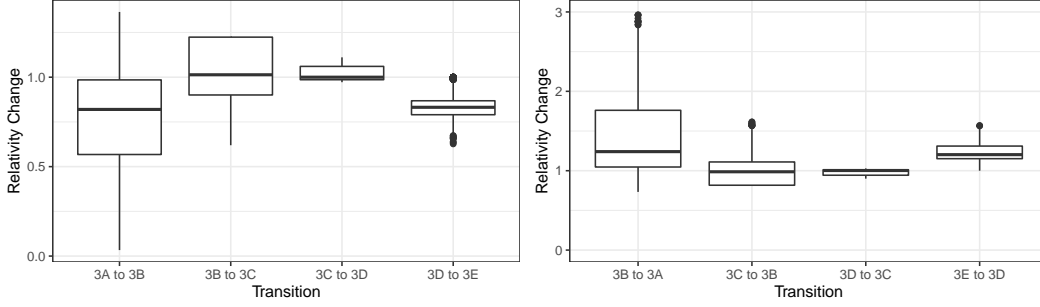
To see the potential impact of such a transition, we used the whole portfolio, and analyzed insureds with a number of items on the boundary of each group 3A, . . . , 3E. Our objective is to measure the impact of a small variation in the number of insured items on the premium calculated by the BMS models. Table 8 shows the 8 possibilities studied. Of course, it would be possible for a farm to add or remove two or more items from its policy in the same year. In those situations, it would be possible for a farm to transition from group 3A to 3C (by adding between 4 and 7 items in the same year), from group 3B to 3D (by adding at least 5 items) and from group 3C to 3E (by adding at least 5 items). Similarly, farm of group 3E can transition to group 3C by removing at least 5 items, and so on. We do not show the impact of those transitions, but they could be analyzed easily following the same procedure.

Nb. of items	Actual Group	New Group if an item is added	Nb. of items	Actual Group	New Group if an item is removed
1	3A	3B	2	3B	3A
4	3B	3C	5	3C	3B
8	3C	3D	9	3D	3C
12	3D	3E	13	3E	3D

*Table 8: List of farms analyzed for the transition study*

For each possibility, we analyze the impact of the transition from one group to another one by computing the ratio of BMS relativities. The result is shown for each transition in Figure 8. Transitioning from group 3A or to group 3A generates extreme situations (the right graph of the figure is limited because the maximum value is around 27). For other transitions, even if the impact is not as big, we can see that changing groups could have a serious impact. Receiving a new past-claims surcharge without making a new claim is not something that could be applied in practice.

Creating distinct BMS models for young and older drivers, as in Denuit et al. (2007), is more logical and simpler. Firstly, because the BMS model to use depends on the age of the driver, which the insured does not control. Secondly, because the transition from the first BMS (for young drivers) to the second BMS only happens once and in one direction. Consequently, the insurer knows when the transition will happen. But, knowing the time of the transition has another advantage. Indeed, the insured's past risk characteristics, i.e.  $X_{i,t}$  for  $t < T$ , are usually



**Figure 8:** *Relativity impact of the transition from one group to another*

not used in ratemaking, and are often not even available to the insurer. If the insurer knows at the age at which the insured moves from the first BMS model to the second one, which we can call  $a$ , it allows the insurer to know that an insured of  $a$  or  $a + 1$  years old has just made a BMS transition. Without knowing the insured’s past characteristics, the insurer knows that it should apply transition rules, which could limit the variation in surcharges or discounts.

If the insurer uses the number of insured items to create distinct BMSs, it is not automatically possible to know if an insured has moved from one BMS to another. It then becomes very difficult to propose a method to mitigate the effect of the transitions, like the ones seen in Figure 8, unless the insurer collects the history of the number of insured items or heavily transforms the BMS pricing model.

## 4 Conclusion

As mentioned in Boucher (2022), using BMS model instead of complex longitudinal distributions or advanced credibility approaches simplifies the understanding of the ratemaking structure of an insurer. Indeed, by its simplicity, the BMS system can be explained to the legislative authorities that regulate pricing, as well as to the various administrators of insurance companies and policyholders. However, despite the quality of prediction of the BMS models, the BMS often appear unfair because it does not consider the initial risk of the insured in the penalty structure.

In order to respond to this criticism of policyholders who might complain about being penalized twice, first by the *a priori* ratemaking, and then by the predictive ratemaking, we proposed to modify the BMS approach. Using a recursive division algorithm such as the one used by Diao and Weng (2019), we created several groups of insureds and applied a separate BMS model for each group. In our numerical application, we used the farm insurance product, and the *a priori* risk was defined as the number of insured items. By using a recursive partitioning algorithm on the farm portfolio, we show that the fit statistics and the predictive scores were greatly improved compared to a unique BMS applied on the whole portfolio. Moreover, we show that partitioning makes it easier to identify groups of insureds that have different loss experience than other groups. Indeed, the approach allows us to see that farms with only one insured item had a very different loss experience than other types of farms.

The approach generates excellent fit statistics and provides very good prediction quality and could easily be applied to many insurance products such as automobile insurance. However, when used with the number of insured items, it has also flaws. The main flaw is related to possible overfitting, but also on the possibility of a transition from one BMS to another. Indeed, because the BMS used for a farm is based on the number of insured items, a farm could quickly move to another BMS simply by adding or removing an item from its policy. As we have seen in our numerical application, this leads to situations where an insured’s premium can rise or fall sharply simply because the applicable structural parameters  $\{\ell_{max}, \ell_{min}, \Psi\}$  change too drastically from one BMS to the next.

A direct way to correct this approach could be to allow the sharing of some structural parameters ( $\ell_{max}$ ,  $\ell_{min}$  and  $\Psi$ ), or some  $\beta$  between each group. For example, in Tables 5 and 7, we saw some similarities between parameters. This approach can save parameters to the model, which often reduces overfitting. This is an interesting idea that could work. However, creating an algorithm to estimate the parameters of such a model could be far more complex than what we used here, and would need to be developed.

We instead believe that it is the step of dividing the portfolio into clear groups that should be generalized. By analogy, using a model that necessitates partitioning of the portfolio is similar to a piecewise regression, where a different regression model is used for each segment of the variable to model. We think that it makes more sense to develop a unique BMS model for all farms where, to follow the analogy, a smoothing function might be used to consider the size, or the *a priori* risk, of each farm. Indeed, because there might be some similarities between contracts or policyholders from different groups, a more holistic approach might be worth considering. This model could then consider the size of the insured as a characteristic of the risk to be included, for example, in the penalty structure of the past-claims algorithm.

## Acknowledgements

The author gratefully acknowledges The Co-operators through the Co-operators Chair in Actuarial Risk Analysis.

## References

- A. Abdallah, J.-P. Boucher, and H. Cossette. Sarmanov family of multivariate distributions for bivariate dynamic claim counts model. *Insurance: Mathematics and Economics*, 68:120–133, 2016.
- K. C. Ágoston and M. Gyetvai. Joint optimization of transition rules and the premium scale in a bonus-malus system. *ASTIN Bulletin: The Journal of the IAA*, 50(3):743–776, 2020.
- L. Bermúdez and D. Karlis. Multivariate inar (1) regression models based on the sarmanov distribution. *Mathematics*, 9(5):505, 2021.
- L. Bermúdez, M. Guillén, and D. Karlis. Allowing for time and cross dependence assumptions between claim counts in ratemaking models. *Insurance: Mathematics and Economics*, 83:161–169, 2018.
- C. Bolancé, M. Denuit, M. Guillén, and P. Lambert. Greatest accuracy credibility with dynamic heterogeneity: the harvey-fernandes model. *Belgian Actuarial Bulletin*, 7(1):14–18, 2007.
- J.-P. Boucher. Bonus-malus scale models: Creating artificial past claims history. *To be published in Annals of Actuarial Science*, 2022.
- J.-P. Boucher and R. Inoussa. A posteriori ratemaking with panel data. *ASTIN Bulletin: The Journal of the IAA*, 44(3):587–612, 2014.
- H. Bühlmann. Experience rating and credibility. *ASTIN Bulletin: The Journal of the IAA*, 4(3):199–207, 1967.
- A. C. Cameron and P. K. Trivedi. *Regression analysis of count data*, volume 53. Cambridge university press, 2013.
- M. Denuit, X. Maréchal, S. Pitrebois, and J.-F. Walhin. *Actuarial modelling of claim counts: Risk classification, credibility and bonus-malus systems*. John Wiley & Sons, 2007.
- L. Diao and C. Weng. Regression tree credibility model. *North American Actuarial Journal*, 23(2):169–196, 2019.
- N. E. Frangos and S. D. Vrontos. Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance. *ASTIN Bulletin: The Journal of the IAA*, 31(1):1–22, 2001.
- E. Gómez-Déniz and E. Calderín-Ojeda. Multivariate credibility in bonus-malus systems distinguishing between different types of claims. *Risks*, 6(2):34, 2018.
- C. Gourieroux and J. Jasiak. Heterogeneous inar (1) model with application to car insurance. *Insurance: Mathematics and Economics*, 34(2):177–192, 2004.

- H. Jeong and E. A. Valdez. Predictive compound risk models with dependence. *Insurance: Mathematics and Economics*, 94:182–195, 2020.
- J. Lemaire. *Bonus-malus systems in automobile insurance*, volume 19. Springer science & business media, 2012.
- M. Mehmet and Y. Saykan. On a bonus-malus system where the claim frequency distribution is geometric and the claim severity distribution is pareto. *Hacettepe Journal of Mathematics and Statistics*, 34(1):75–81, 2005.
- R. Oh, P. Shi, and J. Y. Ahn. Bonus-malus premiums under the dependent frequency-severity modeling. *Scandinavian Actuarial Journal*, 2020(3):172–195, 2020.
- F. Pechon, M. Denuit, and J. Trufin. Home and motor insurance joined at a household level using multivariate credibility. *Annals of Actuarial Science*, pages 1–33, 2019a.
- F. Pechon, M. Denuit, and J. Trufin. Multivariate modelling of multiple guarantees in motor insurance of a household. *European Actuarial Journal*, 9(2):575–602, 2019b.
- J. Pinquet. Positivity properties of the arfima (0, d, 0) specifications and credibility analysis of frequency risks. *Insurance: Mathematics and Economics*, 95:159–165, 2020.
- V. Roel, K. Antonio, and G. Claeskens. Unraveling the predictive power of telematics data in car insurance pricing. *Available at SSRN 2872112*, 2017.
- P. Shi and E. A. Valdez. Longitudinal modeling of insurance claim counts using jitters. *Scandinavian Actuarial Journal*, 2014(2):159–179, 2014.
- P. Shi, X. Feng, and J.-P. Boucher. Multilevel modeling of insurance claims using copulas. *The Annals of Applied Statistics*, 10(2):834–863, 2016.
- R. M. Verschuren. Predictive claim scores for dynamic multi-product risk classification in insurance. *ASTIN Bulletin: The Journal of the IAA*, 51(1):1–25, 2021.