RESEARCH ARTICLE

Construction of rating systems using global sensitivity analysis: A numerical investigation

Arianna Vallarino[1], Giovanni Rabitti[2] and Amir Khorrami Chokami[3]

[1]Collegio Carlo Alberto, Turin, Italy, [2]Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, UK Maxwell Institute for Mathematical Sciences, Edinburgh EH9 3FD, UK and [3]ESOMAS Department, University of Turin and Collegio Carlo Alberto, Turin, Italy
Corresponding author: Giovanni Rabitti; Email: g.rabitti@hw.ac.uk

Abstract

The ratemaking process is a key issue in insurance pricing. It consists in pooling together policyholders with similar risk profiles into rating classes and assigning the same premium for policyholders in the same class. In actuarial practice, rating systems are typically not based on all risk factors but rather only some of factors are selected to construct the rating classes. The objective of this study is to investigate the selection of risk factors in order to construct rating classes that exhibit maximum internal homogeneity. For this selection, we adopt the Shapley effects from global sensitivity analysis. While these sensitivity indices are used for model interpretability, we apply them to construct rating classes. We provide a new strategy to estimate them, and we connect them to the intra-class variability and heterogeneity of the rating classes. To verify the appropriateness of our procedure, we introduce a measure of heterogeneity specifically designed to compare rating systems with a different number of classes. Using a well-known car insurance dataset, we show that the rating system constructed with the Shapley effects is the one minimizing this heterogeneity measure.

1. Introduction

The process of determining premiums for policyholders based on their risk characteristics is known as ratemaking. Ratemaking involves the utilization of rating factors, which divide the insured population into distinct risk classes. While all available risk factors could theoretically be employed, practical considerations may lead to the exclusion of certain factors due to excessively high premiums or to the complexity of the tariff. More precisely, by aggregating risk classes, new classes emerge where some policyholders pay more than their individual premium, while others pay less. This construction ensures that policyholders with excessively high individual premiums pay a reduced tariff premium, thanks to the solidarity of policyholders who would have lower individual premiums in the original granular classes (see Olivieri and Pitacco, 2015 for a comprehensive discussion). Concerning the choice of the rating factors, Kudryavtsev (2009) writes that *insurer portfolios are not homogeneous [...] Actuaries try to identify risk factors (covariates) that help explaining differences between objects insured. Then, the net premium rates are estimated on the basis of conditional loss distributions [...] This approach actually breaks the portfolio up into sub-portfolios [...] Each sub-portfolio is usually more homogeneous than the total insurance portfolio* (p. 297). Then, the challenge lies in determining the "appropriate" risk factors and establishing their connection to the "homogeneity of sub-portfolios" (rating classes). This is the primary focus of the current study. Specifically, we aim to identify the key risk factors that significantly influence insurance premiums, as this insight is valuable for enhancing the interpretability of the pricing model. Subsequently, we construct the rating system based on these identified risk factors and examine whether this selection mitigates heterogeneity within the rating classes.

The starting point of our analysis is the estimation of the individual premiums for all the policyholders in the portfolio by using a regression model with all available risk factors (the complete regression model henceforth). Providing a new methodology to construct individual premiums is not the focus of our work. For this reason, we adopt two recent approaches proposed by Baione and Biancalana (2019) and (2021). These authors consider the so-called Two-Part models that are based on a logistic regression to estimate the claim probability and a subsequent quantile regression (QR) for the claim size. After estimating the individual premiums using both approaches, the next step is to develop a rating system. As mentioned by Kudryavtsev (2009) and Olivieri and Pitacco (2015), it is necessary to select specific risk factors in order to construct the rating system. The main question is then how to find the risk factors that cause the highest variations of the individual premiums. A traditional approach to choose such risk factors is based on the *p*-values relative to the logistic regression coefficients (Heras *et al.*, 2018). Nevertheless, the significance of regression coefficients alone does not provide a comprehensive understanding of the ranking of risk factors in terms of importance, which is a critical aspect to explore. Moreover, *p*-values are not reliable in presence of correlations among the risk factors, as typically happens in insurance dataset. Another complication arises when using *p*-values in Two-Part models. As Baione and Biancalana (2021) point out, *a non-statistically significant p-value of the same regression parameter in the GLM Logit does not imply a non-statistical significance in the QRs*. Thus, we need to adopt a method to find the importance of risk factors in individual pricing with the Two-Part models, which is not based on *p*-values.

The Shapley value, originating from the economic literature (Shapley, 1953), is an attribution method that has found various applications in the actuarial field. It is a versatile tool for allocating the quantity produced by a team among individual players. In the actuarial literature, Lemaire (1984) applies the Shapley value to provide a solution for the problem of operating cost allocation among insurance companies or among different classes of a company. Lemaire (1991) considers other insurance applications of cooperative game theory. Kaye (2005) provides an overview of risk measurement and capital allocation techniques, including the Shapley value. Recently, Godin *et al.* (2023)quantify the contributions of various risk sources to variable annuities contracts using the Shapley value. Mayer *et al.* (2023) provide an overview on approaches based on the Shapley value for explaining machine learning models.

In parallel, the Shapley value has gained increasing attention in the sensitivity analysis literature as a global variable importance index for complex models. Researchers have proposed using the Shapley value to allocate the variance of a model's output among its input variables (Owen, 2014). The underlying idea is that variables explaining the highest fraction of the variance are considered the most important, as their variations have a significant impact on the model's output. Allocating the variance to individual risk factors provides a natural approach to defining global variable importance. Owen suggests using the Sobol' indices as the value function for computing Shapley values (Sobol', 1993). When the Sobol' indices are used as the value function, the resulting Shapley values are referred to as Shapley effects (Song *et al.*, 2016). Shapley effects offer an appealing method for attributing the model's variance to individual risk factors, thereby providing insights into their relative importance. Shapley effects remain defined and interpretable even in the case of dependent risk factors and/or domain irregularities (Owen and Prieur, 2017). To the best of our knowledge, Rabitti and Borgonovo (2020) are the only work in the actuarial literature in which Shapley effects are adopted to find the most important risk between mortality and interest rate in annuity models.

In this paper, we consider the Shapley effects to construct the rating system. Our rationale is that constructing the rating system based on the most important variables serves to minimize the variability within the rating classes, as it reduces the residual source of variance as much as possible. We provide an analytical example confirming the appropriateness of our approach. However, in general Shapley effects cannot be expressed in closed form, even for very simple bivariate models (see Owen and Prieur, 2017). Hence, numerical estimation becomes the only possible approach. For this reason, in this work we combine the algorithm of Plischke *et al.* (2021) (which reduces the computational cost of Song *et al.*, 2016) with the cohort method of Mase *et al.* (2019). To check the validity of our approach in reducing the

intra-class variability, we consider the homogeneity indices of Henckaerts *et al.* (2018). Additionally, we introduce an actuarial indicator that specifically focuses on comparing rating systems. In our case study, all the indicators confirm that the rating system constructed with variables with the highest Shapley effects shows the highest homogeneity within the rating classes. In general, our approach provides also a novel application of the Shapley effects which, beyond traditional model interpretability, might support the rating classes construction.

The rest of the paper is organized as follows: Section 2 provides an overview of the two methods for the individual insurance premium estimation. Section 3 presents a brief description of Shapley effects in the context of variance-based sensitivity measures. In Section 4, we discuss how to construct the rating systems based on Shapley effects, and we introduce the homogeneity measures. Section 5 contains the numerical application on an Australian car insurance dataset.

2. Individual premium estimation

A commonly used approach for individual ratemaking is based on the QR, which was introduced by Koenker and Bassett (1978). The QR allows for the estimation of the relationship between a specific quantile of the random loss of the i-th policyholder and a set of covariates $\mathbf{x}_i = (x_{i1}, \ldots, x_{ik})$.

The application of the QR in insurance ratemaking was first proposed by Kudryavtsev (2009). In this approach, the estimation of the aggregate claims amount is divided into two parts: estimating the conditional aggregate claims amount given that the policyholder has made a claim and assessing the probability of having no claims. However, a limitation of this approach is that it assumes a single probability of having no claims for all risk classes, disregarding the varying levels of riskiness among different classes. Each risk class may have its own distinct probability of having no claims based on the specific characteristics of the policyholders within that class. Ignoring this heterogeneity may result in inadequate pricing decisions and inaccurate estimations of the aggregate claims amount.

To overcome this limitation, Heras *et al.* (2018) extend Kudryavtsev's model through the estimation of a different probability of having no claims for each risk class. Firstly, they adopt a logistic regression to estimate the class-specific probabilities of having no claims. In the second stage, the QR is applied to model the conditional loss distribution given that a claim occurs. Despite that the Two-Part model of Heras *et al.* (2018) offers a more accurate approach, it is computationally inefficient for portfolios with a medium or large number of policyholders, since it requires to run a QR for each risk class. To solve this issue, Baione and Biancalana (2019) consider the specification of a unique probability level for the conditional aggregate claims amount. This approach allows for the calibration of a single QR model. In a subsequent work, Baione and Biancalana (2021) modify the construction of Heras *et al.* (2018) and introduce the parametric quantile regression (PQR) model of Frumento and Bottai (2016) for estimating the conditional aggregate claims amount. In the PQR, regression coefficients are modeled as parametric functions of the order of the quantile, allowing for the estimation of individual insurance premiums with the computation of a single regression model.

From now on, we adopt the following notation. Moreover, let:

- N_i be a binary variable that indicates whether the i-th policyholder has at least one claim, where $i = 1, \ldots, n$ indexes the policies;
- Y_i be the variable describing the total amount of claims for the i-th policyholder (which can be null for those with no claims);

In the first stage of the model, the estimation of the probability of having no claims $p_i^* := \mathbb{P}(N_i = 0 | \mathbf{x}_i)$ is performed via logistic regression. In their works, Heras *et al.* (2018), Baione and Biancalana (2019), and (2021) observe that some policies do not cover the entire year, but only a fraction of it. For this reason, they consider the adjusted probability of having no claims $p_i = p_i^* / \text{exposure}_i$, where exposure_i is the fraction of the year during which the i-th policyholder has been observed. Adjusting for exposure

allows for a fair comparison of claim frequencies across different policyholders or groups with varying levels of exposure. We firstly find the probability p_i^* using the logistic regression model

$$\ln\left(\frac{1-p_i^*}{p_i^*}\right) = \mathbf{x}_i^T \beta, \tag{2.1}$$

then we compute the adjusted p_i (see de Jong and Heller, 2008). For the second stage, we aim to compute $Q_{\theta_i}[Y_i|\mathbf{x}_i]$, that is, the θ_i-quantile of Y_i as in Baione and Biancalana (2019), for simplicity denoted in the following $Q_{\theta_i}[Y_i]$. To begin with, the cumulative distribution function of Y_i is

$$F_{Y_i}(Q_{\theta_i}[Y_i]) = p_i + (1-p_i)F_{Y_i|Y_i>0}(Q_{\theta_i}[Y_i]). \tag{2.2}$$

In Equation (2.2), the cumulative probability of the *i*-th policy not exceeding the total claim amount Y_i is obtained by adding the two terms. The first term represents the probability that the policy has no claims (p_i), while the second term represents the probability that the policy does not exceed Y_i given that it has at least one claim, multiplied by the complementary probability of having no claims ($1-p_i$).

It is typical to set a unique probability level θ for all policyholders (Heras *et al.*, 2018; Baione and Biancalana, 2021). This practice ensures fairness and avoids any potential discrimination among policyholders when assessing individual losses. Hence, Equation (2.2) becomes $\theta = p_i + (1-p_i)\theta_i^*$ and the probability level of the conditional distribution of Y_i given that $Y_i > 0$ is

$$\theta_i^* = \frac{\theta - p_i}{1 - p_i}. \tag{2.3}$$

Then,

$$Q_{\theta}[Y_i] = F_{Y_i|Y_i>0}^{-1}\left(\frac{\theta - p_i}{1 - p_i}\right) = Q_{\theta_i^*}[Y_i \mid Y_i > 0] = \mathbf{x}_i^T \beta_{\theta_i^*}. \tag{2.4}$$

The θ-quantile of the total claims amount Y_i equates the θ_i^*-quantile of the conditional total claims amount Y_i given that $Y_i > 0$. Differently to θ, the probability level θ_i^* depends on i since it is relative to an individual quantile compensation with respect to the level θ. Thus, it is possible to estimate the θ-quantile of Y_i by simply modeling the conditional distribution of the losses given that they are positive. As done in Heras *et al.* (2018) and Baione and Biancalana (2021), we select $\theta = 95\%$ as the probability level of interest (further details are provided in Baione and Biancalana, 2021). We remark that the notation $\beta_{\theta_i^*}$ pertains to the coefficients of the QR, whereas the notation β is associated with those of the logistic regression in Equation (2.1).

At this point, it is necessary to estimate a QR for each θ_i^*, $i = 1, 2, \ldots, n$. Clearly, to estimate individual insurance premiums becomes rapidly unfeasible as the dimension of the portfolio increases. Next Sections 2.1 and 2.2 present two procedures solving this issue.

2.1. Parametric quantile regression ratemaking (PQRR)

Baione and Biancalana (2021) apply in the second stage the PQR to obtain a parsimonious model for the individual premium calculation. With PQR, the regression coefficients are modeled as parametric functions of the order of the quantile, depending on a set of parameters γ

$$\beta_l(\theta, \gamma_l) = \gamma_{l0} + \gamma_{l1}b_1(\theta) + \cdots + \gamma_{ld}b_d(\theta). \tag{2.5}$$

By convention, $b_0(\theta) = 1$ and $l = 0, \ldots, s$. The quantity s represents the total count of numerical variables and all possible levels of the categorical risk factors, excluding the corresponding baseline level for each factor. On the other hand, d denotes the number of functions utilized for performing the PQR analysis. To choose the functions $b_1(\theta), \ldots, b_d(\theta)$, the only assumption to take into account is that their computation must lead to a θ-monotonically increasing quantile function. Baione and Biancalana (2021) suggest to use the shifted Legendre polynomials (SLPs). A polynomial of degree d is defined as

$\widetilde{P}_d(x) = P_d(2x - 1)$ with $\widetilde{P}_d$ orthogonal on [0,1]. The R package `iqr`, used to estimate the PQR, is optimized for this kind of orthogonal polynomials. The choice of the polynomial degree d shall be made in accordance with the dataset. In fact, QR can be affected by quantile crossing problems. This problem appears when quantile curves cross each other and this leads to a nonmonotonic behavior. In the context of individual premium estimation, this issue can be overcome by lowering the degree of the SLP. In fact, even if using higher degrees could appear better since leads to a more flexible model, a quantile crossing issue can arise due to the oscillations of polynomials. For a complete description of this method, we refer to Frumento and Bottai (2016) and Baione and Biancalana (2021). The advantage of this approach is that by calibrating only one regression we are able to model the conditional claims amount at the proper probability level in Equation (2.3) for each policyholder in the portfolio. To compute the insurance premium, Heras *et al.* (2018) and Baione and Biancalana (2021) use a quantile premium principle that is a linear convex combination of the expected value of Y_i and its θ-quantile

$$P_i^{PQRR} = (1 - \alpha)\mathbb{E}[Y_i] + \alpha Q_\theta[Y_i], \tag{2.6}$$

for $\alpha \in [0, 1]$, where we recall that Y_i is a function of the risk factors $\mathbf{x}_i$. The rationale of this premium principle is to charge the fair value of a contract (the expected loss) with a safety loading component, which makes the premium more expensive than the mean. The unconditional expectation $\mathbb{E}[Y_i]$ can be computed using the law of total expectation:

$$\begin{aligned}
\mathbb{E}[Y_i] &= \mathbb{E}\left[\mathbb{E}[Y_i|N_i]\right] \\
&= \mathbb{P}(N_i = 0)\,\mathbb{E}[Y_i|N_i = 0] + \mathbb{P}(N_i > 0)\,\mathbb{E}[Y_i|N_i > 0] \\
&= (1 - p_i)\mathbb{E}[Y_i \mid Y_i > 0],
\end{aligned} \tag{2.7}$$

and the individual premium P_i^{PQRR} is then found with the PQR and the premium principle in Equation (2.6). The corresponding individual premium is denoted with PQRR. For its computation, it is necessary to calibrate the proper level for the loading factor α. The idea is that the sum of all premiums has to equate to the aggregate premium $\pi(Y)$, where Y denotes the random aggregate loss. Therefore, we need to find the value of α that satisfies the following balance equation:

$$\pi(Y) = \sum_{i=1}^{n} [\alpha Q_\theta[Y_i] + (1 - \alpha)\mathbb{E}[Y_i]], \tag{2.8}$$

which yields

$$\alpha = \frac{\pi(Y) - \mathbb{E}[Y]}{Q_\theta[Y] - \mathbb{E}[Y]}. \tag{2.9}$$

Following Baione and Biancalana (2019, 2021), the aggregate premium $\pi(Y)$ has to be computed as the sum of the expected aggregate loss plus a risk margin. The risk margin is selected using the solvency conditions:

$$\pi(Y) = (1 + 3\sigma)\mathbb{E}[Y]. \tag{2.10}$$

The volatility factor σ is set at the level of the 10% from Solvency II Standard Formula for Premium Risk in the Motor Third Party Liability insurance (EIOPA, 2009). In this way, we have indirectly found the coefficient α in a way that is compatible with the Solvency II regulation. We remark that α can be specified directly by the analyst in Equation (2.6). However, Baione and Biancalana (2021) consider this choice too important to be left at a discretionary judgment.

2.2. Quantile regression premium calculation (QRPC)

In the actuarial literature, Baione and Biancalana (2019) propose another method to estimate individual premiums using the QR.

In the second stage of the Two-Part model, a unique probability level is fixed for the conditional total claim amount Y_i given that $Y_i > 0$, in order to have $\theta_i^* \equiv \theta^*$. As a consequence, the probability level for

the total claim amount is no longer unique, but it is different for each policyholder, depending on p_i and θ^*. So, it is possible to run a unique QR using the probability level θ^* for Y_i whenever it is positive. With this intuition, they propose the following premium principle:

$$P_i^{QRPC} = (1 - p_i)Q_{\theta^*}[Y_i \mid Y_i > 0]. \tag{2.11}$$

To calibrate the unique probability level θ^* for all policyholders, Baione and Biancalana (2019) suggest finding the value of θ such that all individual premiums sum up to the aggregate premium $\pi(Y)$

$$\pi(Y) = \sum_{i=1}^{n} P_i^{QRPC} = \sum_{i=1}^{n} (1 - p_i)Q_{\theta}[Y_i \mid Y_i > 0], \tag{2.12}$$

where $\pi(Y)$ is the aggregate premium calculated via Equation (2.10). In particular, θ^* can be found by optimizing the value of θ in Equation (2.12).

Using the QRPC and PQRR models, we find two distinct sets of individual insurance premiums since we use different premium principles and different methods to assess the probability level of interest. In our application in Section 5 we show that, despite differences in premiums at the individual level, at the global level the importance risk factors that influence premiums are the same.

3. From Sobol' indices to Shapley effects

Sensitivity analysis denotes the set of techniques and methods aimed at finding the most influential risk factors driving the output of a scientific model (Saltelli *et al.*, 2008; Borgonovo and Plischke, 2016; Da Veiga *et al.*, 2021). Since this information enables the interpretability of the model (Saltelli *et al.*, 2008), sensitivity analysis of a decision-support model clearly provides essential insights to analysts and supports model-based policy decisions (Borgonovo *et al.*, 2022). In the sensitivity analysis literature, the importance of risk factors is typically quantified partitioning the model output variance into individual risk factor contributions and their interactions. The resulting variance-based sensitivity measures are known as Sobol' indices (Sobol', 1993). The Sobol' indices quantify the variance explained by individual risk factors as well as their interaction effects. The risk factors explaining a higher fraction of the variance are the most important. The Sobol' indices are constructed from the functional ANOVA decomposition of the model output. Let $W = h(\mathbf{X})$ be the input–output model, with the assumptions that $h(\mathbf{X}) \in L^2$ and that the risk factors $\mathbf{X} = (X_1, X_2, \ldots, X_k)$ are independent. The ANOVA decomposition consists in writing $h(\mathbf{X})$ as sum of functions of increasing dimensions:

$$h(\mathbf{X}) = \sum_{u \subseteq K} h_u(\mathbf{X}_u), \tag{3.1}$$

where $\mathbf{X}_u$ contains the coordinates of $\mathbf{X}$ indexed by $u \subseteq K = \{1, \ldots, k\}$. The functions $h_u(\mathbf{X}_u)$ are recursively defined as:

$$h_u(\mathbf{X}_u) = \int \left(h(\mathbf{X}) - \sum_{v \subset u} h_v(\mathbf{X}_v) \right) dB(\mathbf{X}_{-u}) \tag{3.2}$$

where $B(\mathbf{X}_{-u})$ denotes the distribution of the complementary risk factors $-u = \{1, 2, \ldots, k\} \setminus u$ not indexed by u. These functions h_u are orthogonal since $\int h_u(\mathbf{X}_u)h_v(\mathbf{X}_v)dB(\mathbf{X}) = 0$ for $u \neq v$.

The intuition for quantifying the importance of the group $\mathbf{X}_u$ is evaluating how much variance of W it explains. Denoting with σ_u^2 the variance $\sigma_u^2 = \mathrm{Var}(h_u(\mathbf{X}_u))$, by orthogonality it holds

$$\mathrm{Var}(W) \equiv \sigma^2 = \sum_{u \subseteq K} \sigma_u^2. \tag{3.3}$$

In the framework of global sensitivity analysis, Sobol' (1993) introduces the importance index

$$\underline{\tau}_u^2 = \sum_{v \subseteq u} \sigma_v^2. \tag{3.4}$$

Since $\underline{\tau}_u^2 = \text{Var}(\mathbb{E}[W|\mathbf{X}_u]) = \text{Var}(W) - \mathbb{E}[\text{Var}(W|\mathbf{X}_u)]$, this index can be interpreted as the expected reduction of the variance of the output W that is achieved when the set of u risk factors is fixed. In the computer experiment literature, the Sobol' indices are computed using the pick-and-freeze sampling strategy (see Saltelli *et al*., 2008). Sobol' indices are interpretable in case of independent risk factors, but it is no longer true when the independence assumption is violated. To overcome the problems with dependent variables and domain irregularities, a promising solution is the adoption of Shapley value for the variance allocation instead of Sobol' indices (Owen and Prieur, 2017).

3.1. Shapley value

The Shapley value (Shapley, 1953) is an attribution method originated from game theory. It is used to allocate the total value generated by a coalition of players into the contribution of each individual player in a coalition.

Denote with ϕ_j the Shapley value of the j-th player. Consider a set of players $K = \{1, 2, \ldots, k\}$ that participate in a game. We denote with $u \subseteq K$ any possible subcoalition of players and with val $: 2^K \to \mathbb{R}$ the value function of the game, with $\text{val}(\emptyset) = 0$. The value generated by the players in the subset u is $\text{val}(u)$ and the total value of the game to be allocated is $\text{val}(K)$.

Consider the following desirable properties for an attribution method (Winter, 2002):

- *Efficiency:* The sum of Shapley values of all players equates to the total value of the game $\sum_{j=1}^k \phi_j = \text{val}(K)$. This means that the total payoff of the game is allocated among all players.
- *Symmetry:* If players j_1 and j_2 make the same marginal contribution to any coalition, $\text{val}(u \cup \{j_1\}) = \text{val}(u \cup \{j_2\})$ for all $u \subseteq K$ with $j_1, j_2 \notin K$, then $\phi_{j_1} = \phi_{j_2}$. Thanks to this axiom, the order in which players are considered does not influence the value of the coalition; players with the same marginal contribution receive the same payoff.
- *Dummy:* If j is a dummy player, $\text{val}(u \cup \{j\}) = \text{val}(u)$ for all $u \subseteq K$, then $\phi_j = 0$. In other words, if the player's contribution to the coalition is zero, the payoff of the player is zero (axiom of the null-player).
- *Additivity:* If val and val$'$ have Shapley values ϕ and ϕ', respectively, then the game with value $\text{val}(u) + \text{val}'(u)$ has Shapley value $\phi_j + \phi'_j$ for all $j \in K$. This means that if two games are combined, the payoff assigned to the j-th player is equal to the sum of the two payoffs he would get playing the two games separately.

For the generic player j, Shapley (1953) proves that the unique value satisfying these four properties can be expressed as:

$$\phi_j = \sum_{u \subseteq K \setminus \{j\}} \frac{(k - |u| - 1)! |u|!}{k!} (\text{val}(u \cup \{j\}) - \text{val}(u)) \tag{3.5}$$

This index is based on the marginal increase $\text{val}(u \cup \{j\}) - \text{val}(u)$ generated when the j-th player joins the coalition u. The marginal contribution is then averaged over all possible coalitions.

The Shapley value is attracting increasing attention in the sensitivity analysis literature as a measure of variable importance (Owen, 2014; Song *et al*., 2016). Owen (2014) suggests to use as value function the Sobol' index

$$\text{val}(u) = \underline{\tau}_u^2. \tag{3.6}$$

The Shapley values obtained with this value function are called the *Shapley effects* (Song *et al*., 2016). Shapley effects enjoy appealing properties:

1. $\phi_j \geq 0$ for all $j = 1, 2, \ldots, k$. This holds by Equation (3.5) and because $\underline{\tau}_{u \cup \{j\}}^2 - \underline{\tau}_u^2$ is positive as the variance explained increases when a new risk factor is considered for all $u \subseteq K$.
2. $\sum_{j=1}^k \phi_j = \text{Var}(W)$. This follows from the efficiency property, meaning that we allocate the variance of the model output.

Remarkably, Properties 1 and 2 hold regardless the dependence structure and marginal distributions of the risk factors. This feature contributes significantly to the widespread adoption of Shapley effects as sensitivity measures in the interpretation of complex models.[1]

The computation of the Shapley effects is not a trivial issue. Song *et al.* (2016) propose a double loop Monte Carlo for the estimation of the value function to be inserted in the Shapley representation in Equation (3.5). Plischke *et al.* (2021) propose an algorithm based on the concept of Möbius inverses, which reduce the number of value functions to be evaluated with respect to the approach of Song *et al.* (2016). With this algorithm, the number of value functions to be evaluated is remarkably reduced from $k! \cdot k$ to $2^k - 1$. Furthermore, it should be noted that the Plischke algorithm offers a way to estimate Shapley values for interactions. However, since it goes beyond the scope of this work, a thorough exploration of this topic is not included here. For more detailed insights, interested readers can refer to the works of Plischke *et al.* (2021) and Lundberg *et al.* (2020).

Denote with $\mathrm{mob}(u)$ the Möbius inverses of the value function $\mathrm{val}(u)$ with:

$$\mathrm{val}(u) = \sum_{v \subseteq u} \mathrm{mob}(v) \quad \text{and} \quad \mathrm{mob}(u) = \sum_{v \subseteq u} (-1)^{|u|+|v|} \mathrm{val}(v). \tag{3.7}$$

Note that, in the case of independent risk factors, we find $\mathrm{mob}(u) = \sigma_u^2$. Equivalently to the representation in Equation (3.5), Shapley values can be written as

$$\phi_j = \sum_{u:j \in u} \frac{\mathrm{mob}(u)}{|u|}. \tag{3.8}$$

Following Plischke *et al.* (2021), let all $2^k - 1$ non-vanishing subsets be indexed by u_m with $m = 1, \ldots, 2^k - 1$. The value functions are included in a vector H with coordinates $H_m = \mathrm{val}(u_m)$. To obtain subset inclusion, a matrix Z is proposed:

$$Z_{mf} = \begin{cases} 1 & \text{if } u_m \subseteq u_f \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

Then, Möbius inverses are obtained with $M = HZ^{-1}$ and Shapley effects are given by

$$\phi_j = \sum_{m:j \in u_m} \frac{M_m}{|u_m|}. \tag{3.10}$$

Now, the Shapley effects computation requires an estimate of the value function in Equation (3.6) to compute the Möbius inverses. Differently from the computer experiment literature, we need a method to estimate the Sobol' indices from given data. A solution to this issue comes from the recent work of Mase *et al.* (2019), who propose the estimation of $\mathrm{Var}(\mathbb{E}[W|\mathbf{X}_u])$ using a cohort approach, as presented in the next section.

3.2. The cohort approach

Assume we have available at hand a dataset of observations $\{(w_i, \mathbf{x}_i)\}_{i=1}^n$, where w_i denotes the quantity of interest. The cohorts are subsets of observations sharing the values of risk factors similar to the ones of a target observation denoted by t. Let x_{ij} be the value of the j-th risk factor for the i-th observation (with $j = 1, \ldots, k$ and $i = 1, \ldots, n$) and let x_{tj} be the target value for the selected risk factor j. It is necessary to construct a similarity function z_{tj} for each risk factor j. In particular, $z_{tj}: X_j \to \{0, 1\}$. The similarity function is equal to 1 when the subject i is considered similar to the target subject t for the risk factor j.

---

[1]The quantile pricing models presented in Section 2 can be considered complex models since they are based on the product of two GLMs (one for the claim sized and one for the claim frequency). Multiplying together two linear models generates non-linear effects among the risk factors.

In the case of real-valued covariates, it can be appropriate to define similarity using a measure of relative distance. A possible similarity function is

$$z_{tj}(x_{ij}) = \begin{cases} 1 & |x_{ij} - x_{tj}| \leq \delta_j, \\ 0 & \text{otherwise}, \end{cases} \tag{3.11}$$

where δ_j is the specified distance. If the variables are continuous, Mase *et al.* (2019) consider $\delta_j = 0.1 \cdot \text{range}(x)$, while $\delta_j = 0$ if they are discrete. Now it is possible to define the cohort of t with respect to risk factors in u as

$$C_{t,u} = \{i \in 1 : n \quad | \quad z_{tj}(x_{ij}) = 1, \forall j \in u\}. \tag{3.12}$$

By definition, $C_{t,u}$ is the set of subjects that are similar to the target subject t for all risk factors $j \in u$. By construction, $C_{t,u}$ is never empty since it contains at least the target observation x_t. Using the cohorts, Mase *et al.* (2019) found estimates of the Sobol' indices as follows. Consider the cohort means

$$\frac{1}{|C_{t,u}|} \sum_{i \in C_{t,u}} w_i = \bar{w}_{t,u}, \tag{3.13}$$

where $|C_{t,u}|$ is the cardinality. It holds that the cohort average $\bar{w}_{t,u}$ is an estimate of the conditional mean for the target $\mathbb{E}[W|X_{t,u}]$. Since each observation has probability n^{-1}, we obtain an estimate of the value function:

$$\text{val}(u) = \frac{1}{n} \sum_{t=1}^{n} \left(\bar{w}_{t,u} - \bar{w} \right)^2, \tag{3.14}$$

where $\bar{w}$ is an estimate of the model output mean $\mathbb{E}[W]$. Hence, to compute the Shapley effects, we use the value function in Equation (3.14), estimated via cohort approach, to calculate the Möbius inverses in Equation (3.7). Finally, Shapley effects are obtained via Equation (3.8), adopting the algorithm proposed by Plischke *et al.* (2021). A MATLAB implementation can be found at https://github.com/giovanni-rabitti/ratingsystemshapley.

4. Constructing and comparing rating systems

After the implementation of the methods presented in Sections 2 and 3, we have at hand the individual quantile premiums (found with QRPC and PQRR) for the $i = 1, 2, \ldots, n$ policyholders in the portfolio. Additionally, we estimate the most important risk factors determining these premiums. This information is crucial for the construction of a rating system in which the policyholders are divided into risk classes as homogeneous as possible.

Suppose that an actuarial analyst has evaluated the Shapley effects for the k risk factors and that this analyst wants to construct a system based on two of them. The two risk factors with highest Shapley effects explain together the highest fraction of the variance of the individual premiums among all possible combinations of two risk factors out of k. Denote them with ϕ_{k_1} and ϕ_{k_2}. Then, $\left(\phi_{k_1} + \phi_{k_2} \right) / \sigma^2$ is the maximum possible. By the efficiency property, we have that $\left(\sum_{j=1}^{k} \phi_j \right) / \sigma^2 = 1$, implying that

$$\frac{\sum_{k_1 \neq j \neq k_2} \phi_j}{\sigma^2} = 1 - \frac{\phi_{k_1} + \phi_{k_2}}{\sigma^2} \tag{4.1}$$

is the minimum possible. In other words, the fraction of variance explained by the remaining $k - 2$ factors is the smallest possible. As a consequence, the variability of the individual premiums within the risk classes based on the risk factors k_1 and k_2 is driven only by the other $k - 2$ factors, and it is the minimum. This implies that in the selected rating system the insurance premium for each policyholder is closer to the average risk class premium than in any other systems.

4.1. Analytical example

We illustrate our procedure with an analytical example. Assume that the generic premium P^* can be written as

$$P^* = \sum_{j=1}^{k} \lambda_j X_j \tag{4.2}$$

where each variable X_j is considered to be discrete (continuous variables are typically discretized for rating systems construction – see Henckaerts *et al.*, for $j = 1, \ldots, k$. Every realization of $\mathbf{X}$ is $\mathbf{x}$, which is the vector of the characteristics of a policyholder (for the sake of simplicity, we drop the index i in this analytical example).

Assume that the rating system is constructed using the selected variables $u \subseteq K$. In this system, a generic class is indexed by r with $r = 1, \ldots, R$ and is uniquely identified by fixing the levels of the vector $\mathbf{X}_u = \mathbf{x}_u$. In this class, a generic premium is given by

$$P_r = \sum_{j \in u} \lambda_j x_j + \sum_{j \notin u} \lambda_j X_j. \tag{4.3}$$

In other words, since the other variables $\mathbf{X}_{-u}$ can vary, P_r is a random variable representing the individual policyholders' premiums in the class r. In particular, when $u = \{1, \ldots, k\}$, then r is the cohort of policyholders sharing the same levels for all variables and P_r becomes deterministic. In the present framework, the following result holds.

Proposition 4.1. *Assume that the class r is identified by $\mathbf{x}_u$. For the any policyholder in r, consider the difference between their individual premium and the mean class premium, $P_r - \mathbb{E}[P_r]$. Then, for any fixed threshold $\xi > 0$, the probability $\mathbb{P}(|P_r - \mathbb{E}[P_r]| \geq \xi)$ is minimized when the rating system is constructed by using the rating factors with the highest Shapley effects.*

Proof. First of all, by the additivity and efficiency axioms of the Shapley values and the independence of the X_is, the variance of P^* is

$$\text{Var}(P^*) = \sum_{j=1}^{k} \phi_j, \tag{4.4}$$

where $\phi_j = \lambda_j^2 \text{Var}(X_j)$ for all $j = 1, \ldots, k$. Analogously, the variance of the individual premiums of policyholders in the class r is

$$\text{Var}(P_r) = \sum_{j \notin u} \phi_j \tag{4.5}$$

because of Equation (4.3). Moreover, the average class premium in r is

$$\mathbb{E}[P_r] = \sum_{j \in u} \lambda_j x_j + \sum_{j \notin u} \lambda_j \mathbb{E}[X_j]. \tag{4.6}$$

This mean value represents the premium that all policyholders in class r are required to pay. Then, it holds

$$\mathbb{P}(|P_r - \mathbb{E}[P_r]| \geq \xi) = \mathbb{P}\left(\left| \sum_{j \notin u} \lambda_j X_j - \sum_{j \notin u} \lambda_j \mathbb{E}[X_j] \right| \geq \xi \right)$$

$$\leq \frac{1}{\xi^2} \text{Var}\left(\sum_{j \notin u} \lambda_j X_j \right) = \frac{1}{\xi^2} \sum_{j \notin u} \phi_j \tag{4.7}$$

by Chebyshev's inequality applied to the random variable $\sum_{j \notin u} \lambda_j X_j$, which represents the residual variability within the system constructed by fixing $\mathbf{X}_u$. If $\mathbf{X}_u$ are the rating factors with the highest Shapley

effects, then the upper bound of $\mathbb{P}(|P_r - \mathbb{E}[P_r]| \geq \xi)$ is the lowest possible, because for any fixed ξ the value $\sum_{j \notin u} \phi_j$ is minimized. Hence, for any class in this rating system, the probability that individual premiums deviate from the class mean is minimized. $\square$

It is worth commenting Proposition 4.1. First of all, Equation (4.5) is a within sum of squares index for the individual premiums in the class r (since we are summing up the residual variances of the individual premiums in the class r). However, differently from clustering methods, our aim is to minimize this index by changing the selected rating factors and not by aggregating factor levels. Secondly, we assumed a linear pricing model with independent and discrete inputs. This choice was necessary to obtain a closed form expression of the Shapley effects (as a matter of fact, Owen and Prieur, 2017 managed to provide analytical expressions of Shapley effects only for some specific bivariate cases).

4.2. Homogeneity indices for system comparison

In general, numerical methods represent the only possible approach to estimate these sensitivity indices. Consequently, it becomes necessary to include statistical indicators to assess the reliability of the our procedure. Consider the premium P_i^*, which is the loading premium computed at the individual level for each policyholder with QRPC or PQRR presented in Equations (2.6) and (2.11). To measure the difference of the individual premiums when computed with all risk factors and with the selected risk factors with the Shapley effects, we propose the mean-squared difference (MSD) defined as

$$\text{MSD} = \frac{1}{n} \sum_{i=1}^{n} \left(P_i^* - P_i^{\text{red}} \right)^2, \tag{4.8}$$

where P_i^{red} represents the class premium of the i-th policyholder computed with the reduced model. The lower the MSD, the closer the individual premium is to the class premium in the rating system. We observe that P_i^{red} coincides with the class premium to which the i-th policyholder belongs to (in every risk class all policyholders pay the same average price), so it is the empirical counterpart of Equation (4.6). To emphasize the dependency on the class r rather than on the policyholder i, we can denote the same average premium as P_r^{red}. Hence, we can equivalently rewrite the MSD as

$$\text{MSD} = \frac{1}{n} \sum_{r=1}^{R} \sum_{i_r=1}^{n_r} \left(P_{i_r}^* - P_r^{\text{red}} \right)^2, \tag{4.9}$$

where n_r is the number of insured in the risk class r, with $\sum_{r=1}^{R} n_r = n$. Equation (4.9) tells us that the MSD is the sum of the deviations of the individual premiums with respect to their class means, averaged for all policyholders in all classes. Hence, the numerator of the MSD is a sum of squared distances, which can be interpreted as a sum of within-classes variations.

We can also finally consider two measures of deviation used in the actuarial literature (Henckaerts *et al.*, 2018), namely the goodness of variance fit (GVF)

$$\text{GVF} = 1 - \frac{\sum_{r=1}^{R} \sum_{i_r=1}^{n_r} \left(P_{i_r}^* - P_r^{\text{red}} \right)^2}{\sum_{i=1}^{n} \left(P_i^* - \overline{P} \right)^2}, \tag{4.10}$$

where $\overline{P} = 1/n \sum_{i=1}^{n} P_i^*$ and the tabular accuracy index (TAI)

$$\text{TAI} = 1 - \frac{\sum_{r=1}^{R} \sum_{i_r=1}^{n_r} |P_{i_r}^* - P_r^{\text{red}}|}{\sum_{i=1}^{n} |P_i^* - \overline{P}|}. \tag{4.11}$$

The denominator in the previous formulas measures the deviation of each individual premium from the global mean. The numerator measures the deviation of each individual premium from its class mean.

The closer the GVF and TAI values are to 1, the more uniform the classes become in terms of their homogeneity. Lastly, note it holds

$$\text{GVF} = 1 - \frac{n}{n-1} \frac{\text{MSD}}{\sum_{j=1}^{k} \phi_j}, \tag{4.12}$$

formally connecting the GVF with the MSD and the Shapley effects. We remark that the GVF provides insights on how close the rating system is with respect to the lowest heterogeneity case.

In practice, the rating classes have different means and heterogeneous variances and they are unbalanced. Thus, the problem of their comparison in terms of internal variability intensifies. In general, when the means of classes differ significantly, comparing the absolute differences in variability (i.e. TAI) or the variances (i.e. MSD) can be misleading. For example, in Equation (4.9) the MSD does not explicitly consider the relative variations within each class. When comparing rating systems, it is important to account for this effect in order to ensure an accurate and effective comparison. This is because the absolute differences may be influenced more by the class means rather than the inherent variability within each class. However, this is not the case when considering measures of relative variability, which account for these differences in means. One of such measures is the coefficient of variation (Olivieri and Pitacco, 2015)

$$\text{CV} = \frac{\sqrt{\sigma_P^2}}{\overline{P}} \tag{4.13}$$

where σ_P^2 is the variance of the premiums. Olivieri and Pitacco (2015) state that the CV is suitable to compare the riskiness of classes composed of a different number of policyholders. Mahmoudvand and Oliveira (2018) use the CV to measure concentration risk in a loan portfolio, while in the actuarial literature Donnelly (2014) adopts it to quantify the mortality risk in a defined-benefit pension scheme.

When comparing rating systems in terms of overall homogeneity of their risk classes, we can adopt the coefficients of variation for every risk class and then average. However, a simple arithmetic average of all coefficients of variations of the risk classes of a rating system might not be a good indicator of the general homogeneity of that system. Namely, a risk class with fewer insured that might be more homogeneous will receive the same weight as a risk class with more insured, producing a distortion. Since different rating systems are based on different combinations of the selected risk factors, they are based on different risk classes. We want to aggregate the coefficient of variations of the risk classes of a rating system in order to make comparisons between rating systems. We consider a linear combination of the coefficients of variation. Precisely, an arithmetic mean would be a misleading indicator of homogeneity since classes with different sizes would receive the same weight. Conversely, adopting weighted means can give the right weight to the homogeneity of risk classes composed of a large number of insured. Hence, we propose to compare rating systems using the weighted mean of coefficient of variations (WMCV)

$$\text{WMCV} = \frac{\sum_{r=1}^{R} n_r \cdot \text{CV}_r}{\sum_{r=1}^{R} n_r}. \tag{4.14}$$

With this approach, we are able to firstly compute the homogeneity of each risk class (with the coefficient of variation) and, secondly, we take the weighted mean to have a measure for the entire rating system. We consider the WMCV the most appropriate index to compare rating systems in terms of overall homogeneity.

5. Numerical application

We perform the steps previously described using a well-known Australian car insurance dataset, available in the R package `InsuranceData`. The dataset represents an insurance portfolio composed of 67,856 policies, of which 4624 have at least one claim. The description of the variables contained in the dataset is provided in Table 1. For a detailed description of the dataset, see de Jong and Heller (2008).

Table 1. *Variables description.*

Variables	Levels/Range
Driver's age band	1 (youngest), 2, 3, 4, 5, 6
Gender	Male and Female
Driver's area of residence	A, B, C, D, E, F
Vehicle age	1(new), 2, 3, 4
Vehicle body type	Bus, convertible, coupe, hatchback, hardtop, motorized caravan – combi, minibus, panel van, roadster, sedan, station wagon, truck, utility vehicle
Vehicle value[†]	0–34.56 (in $10,000s)

[†]Discretized following de Jong and Heller (2008).

The first five risk factors listed in Table 1 are factor variables (whose corresponding levels are mentioned), while *vehicle value* is a continuous variable. As done in de Jong and Heller (2008), we discretize the variable *vehicle value*. This is a necessary step to avoid the number of risk classes exploding in the construction of the rating systems. For this reason, we follow de Jong and Heller (2008) and discretize the risk factor into six different intervals with the same values. We assign to each observation the mean value of the interval where the observation falls. The intervals are the following: $(0, 25, 000\$]$ with 54,971 policies, $(25, 000\$, 50, 000\$]$ with 11,439 policies, $(50, 000\$, 75, 000\$]$ with 1265 policies, $(75, 000\$, 100, 000\$]$ with 104 policies, $(100, 000\$, 125, 000\$]$ with 44 policies, and finally $(125, 000\$, \infty)$ with 33 policies.

The presented variables in Table 1 are the six risk factors that we use in our Two-Part models to estimate the individual premiums. In the dataset, there are four other variables: *Exposure*, *Claim occurrence*, *Claim amount*, and *Number of claims*. *Exposure* is a continuous variable bounded between 0 and 1: the value of one indicates that the policy covers the entire year, while a value less than one means that it covers only the corresponding part of the year. *Claim occurrence* is a binary variable in which 0 is assigned to policies with no claims occurrence and 1 to the ones that suffered at least one claim. These policies also have a positive value of the variable *Claim amount*, which indicates the insurance company's loss. Of course, the policies with zero claims correspond to a claim amount value of zero dollars. The variable *Numbers of claims* counts how many accidents the policyholder has suffered during the year.

All computations have been performed on a standard laptop with 8GB RAM. The PQRR and QRPC have been implemented in the R-software and took around 14 and 2 min, respectively. The estimation of the Shapley effects took around 94 min using the MATLAB software. All codes are available at https://github.com/giovanni-rabitti/ratingsystemshapley.

5.1. Individual premium estimation

The first step of the Two-Part model consists of the estimation of the probability of having no claims. In this case, we use a GLM regression of the binomial family with a logit link function adjusted for the exposure in Equation (2.1). The R code for the logit exposure adjusted is provided by de Jong and Heller (2008). In Table A.1, we report the parameters of the logit GLM regression obtained via the maximum likelihood estimation.

In the second step, we estimate the severity component. We first perform the PQR to find the premiums with the PQRR approach (see Section 2.1). The selected function $b(\theta)$ to implement the procedure is a SLP of degree 1. In particular, the regression coefficient has the following formulation: $\beta_l(\theta, \gamma_l) = \gamma_{l0} + \gamma_{l1}(2\theta - 1), l = 0, \ldots, 27.$[2]

---

[2]Recall that $l = 0$ is the intercept of the regression, while $l = 1, \ldots 27$ are the parameters in the regression for risk factors.

***Table 2.*** *Wald test for rating factors.*

	chi-square	df	*p*-value
(Intercept)	526.569	2	< 2e-16***
veh_age	21.915	6	0.00125**
veh_value	2.053	2	0.35832
veh_body	62.371	24	2.93e-05***
agecat	36.017	10	8.36e-05***
area	31.266	10	0.00053***
gender	5.486	2	0.06438

Baione and Biancalana (2021) use in the same dataset a SLP of degree 3, since in their numerical application the purpose is the estimation of insurance premiums for the eight risk classes that compose the rating system they construct. In our case, we need to use a SLP of degree 1 since a degree equal to 3 produces numerical instability. In particular, using a polynomial of the second or third degree, we obtain some exceptional small/large values (outliers) for the severity component. Figure A.2 in the Appendix A shows that the goodness-of-fit of this model is satisfactory despite the low degree. Estimated parameters of the PQR are reported in Table A.4 in the Appendix A. In Table 2, the Wald test for the significance of rating factors is shown. From Table 2, we can see that the most statistically significant risk factors are *vehicle body*, *agecat*, and *area*. However, as previously explained, a selection based on the *p*-values might be misleading due to the lack of interpretability of the ranking in terms of importance of the risk factors. To solve this critical issue, we use the Shapley effects approach to be able to obtain factor prioritization. The Wald test for the $b(\theta)$ components is also provided in Table A.5 in the Appendix. Using Equation (2.6), we estimate the quantile premium principle. As previously done in Heras *et al*. (2018) and Baione and Biancalana (2021) on this dataset, we use a Gamma GLM with log link function to estimate the conditional expectation of the loss distribution (regression coefficients are provided in Table A.3 in the Appendix).

Finally, we need to calibrate the loading factor α using Equation (2.9). The value obtained through the calibration procedure is $\alpha = 6.84\%$. The severity component can also be estimated using QR and the individual premiums can be found by applying the QRPC approach (explained in Section 2.2): this method requires to find the probability level θ^* such that the balance Equation (2.12) holds. In our application, the resulting value is $\theta^* = 78.60\%$. We recall that we are able to run a unique QR with this probability level. The obtained parameters estimates are shown in Table A.2 (see the Appendix). At this point, we can compute the individual insurance premium using the quantile premium principle in Equation (2.11).

5.2. Application of Shapley effects to the rating system construction

Once individual premiums are estimated, we apply the Shapley effects algorithm to the datasets $\left\{\left(P_i^{QRPC}, \mathbf{x}_i\right)\right\}_{i=1}^n$ and $\left\{\left(P_i^{PQRR}, \mathbf{x}_i\right)\right\}_{i=1}^n$ to identify the two risk factors that, in combination, explain a greater amount of variance compared to any other pair of variables. The utilization of Shapley effects is further justified by the dependence analysis depicted in Figure A.1 in the Appendix. This analysis reveals that the risk factors are not independent, strengthening the rationale for employing the Shapley effects algorithm. Given the presence of six discrete risk factors, it is important to define the notion of similarity between two observations within a subset u. Specifically, two observations are considered similar with respect to subset u if they share the same values for the coordinates indexed by u. In the case of discrete risk factors, the similarity function in Equation (3.11) is set to $\delta_j = 0$ (as done in Mase *et al*., 2019). By applying the algorithm to the premiums obtained with QRPC and PQRR, we display the results in Figure 1. Risk factors *agecat* and *vehicle body* explain together 82.45% of individual premium total variance in case of the QRPC approach and 77.77% in case of the PQRR approach. These two risk
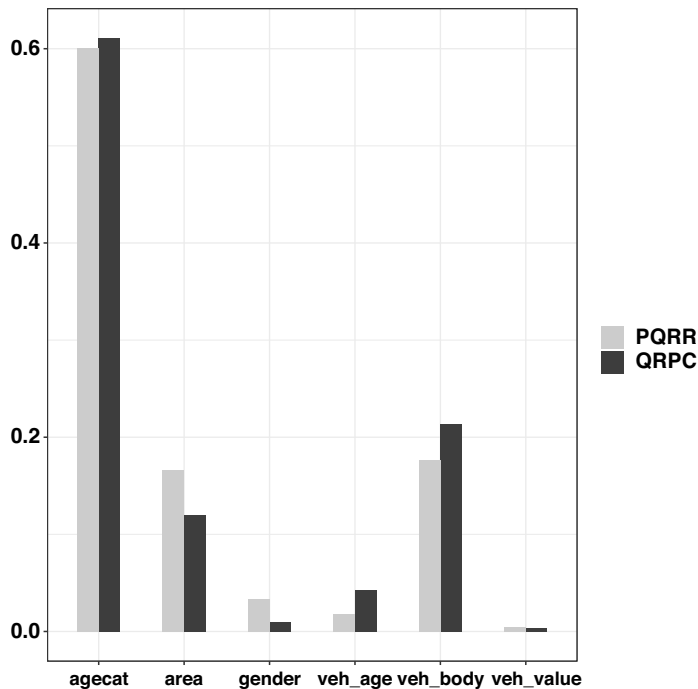
Figure 1. *Shapley effects for the individual premiums computed with QRPC and PQRR. The y-axis shows the Shapley effects for each risk factor (displayed on the x-axis). By the efficiency property, the Shapley effect normalized by the variance can be interpreted as the proportion of explained variance by a single risk factor.*

factors are the greatest drivers of the individual premiums. Conversely, *vehicle value* is negligible and could be "frozen" to a constant value, reducing the model complexity (see, for instance, Saltelli *et al.*, 2008 for a discussion on freezing irrelevant factors). Also, *gender* and *vehicle age* have a low impact on the variability of the individual premiums. These insights enable model interpretability since the analyst knows the main drivers of the pricing models. We note that, despite some differences in premiums at the individual level between the two pricing methods, at the global level the importance of risk factors is the same for both QRPC and PQRR premiums, as shown in Figure 1. Only the fourth and fifth most important risk factors are switched between QRPC and PQRR: *vehicle age* and *gender* are, respectively, the fourth and fifth most important risk factors for QRPC, while their roles are inverted for the PQRR. However, the magnitude of the fourth highest Shapley effect is relatively small (around 3–4% for the two pricing methods). We lastly note that discretizing the variable *veh_value* does not change its irrelevant importance. If it is left continuous, its Shapley effects are 0.0014 for the QRPC and 0.0047 for the PQRR, while as a discrete variable they are 0.0036 and 0.0150, respectively.

5.3. Rating systems comparison

The estimation of Shapley effects provides us insights about the most important risk factors from a statistical point of view. On the other hand, we need an actuarial valuation to demonstrate that the rating system we have constructed is the one with less variability in the insurance premium. We adopt the indices presented in Subsection 4.2 as a measure to assess the variability within each risk class in a rating system and to evaluate the riskiness of the overall rating system.

With our dataset, composed of 6 risk factors, we can construct 15 different systems made up of 2 risk factors each (remember that we refer to the discretized variable *vehicle value*). In the literature,

Table 3. *Weighted mean coefficient of variation, mean-squared difference, and goodness of variance fit.*

Rating system	WMCV		MSD		GVF		Number of risk classes
	PQRR	QRPC	PQRR	QRPC	PQRR	QRPC	
veh_age – gender	0.328	0.365	15673.898	19559.036	0.043	0.031	8
veh_body – veh_age	0.295	0.314	12973.676	15145.669	0.208	0.250	49
veh_body – gender	0.294	0.321	12976.694	16006.447	0.208	0.207	26
veh_body – agecat	**0.139**	**0.143**	**3418.100**	**3839.513**	**0.791**	**0.810**	**78**
veh_age – area	0.301	0.341	13106.800	17166.138	0.200	0.150	24
gender – agecat	0.181	0.211	5519.245	7382.024	0.663	0.634	12
veh_body – area	0.263	0.292	10366.884	13338.936	0.367	0.339	76
veh_age – agecat	0.189	0.197	5985.256	6908.135	0.635	0.659	24
gender – area	0.297	0.348	12897.895	17646.811	0.213	0.126	12
area – agecat	0.152	0.183	3639.166	5674.259	0.778	0.719	36
veh_age – veh_value	0.332	0.366	16058.956	19706.286	0.020	0.024	21
gender – veh_value	0.329	0.372	15833.626	20090.619	0.034	0.005	12
veh_body – veh_value	0.298	0.322	13262.095	16200.813	0.191	0.198	48
agecat – veh_value	0.191	0.213	6146.620	7706.653	0.625	0.618	36
area – veh_value	0.303	0.348	13292.165	17782.359	0.189	0.119	33

two examples of rating systems constructed with this dataset are provided. Heras *et al.* (2018) build a system using the risk factors *vehicle age* and *agecat* selected with the *p*-values resulting from the logistic regression model. However, *vehicle age* is not an important risk factor as Figure 1 shows. The second application is offered in Baione and Biancalana (2019) and (2021): the rating system is constructed by using *gender* and *vehicle age* to simplify the representation of their models.

For every rating system based on two risk factors, we estimate the weighted mean of the variation coefficient via Equation (4.14) and the number of non-empty risk classes. Furthermore, we compute the MSD between the individual premiums using the quantile models with all risk factors included and quantile models based on two risk factors only. Note that, for the latter case of two factors, the PQRR and the QRPC for each rating system were recalibrated.[3]

The results, as presented in Table 3, demonstrate that the rating system built on the factors of *vehicle body* and *agecat* exhibits the lowest overall variability in average premiums within each risk class. Additionally, from Table 3 we note the following remarkable aspects. To begin with, all rating systems constructed using the risk factor *agecat* exhibit low values of WMCV. This was expected since *agecat* explains the largest fraction of the variance of the premiums. The MSD is also minimized for both models, when the risk factors are selected using Shapley effects. Consequently, the GVF is maximized as per Equation (4.12) note that it takes values around 0.8, which are close to 1 (recall that a GVF close 1 indicates a high degree of homogeneity). Furthermore, it is important to point out that the low WMCV and the high GVF are not a consequence of the high number of classes. In fact, the selected rating system with *vehicle body* and *agecat* has 78 risk classes, but also the rating system composed of *vehicle body* and *area*, for instance, has a high number of risk classes (76). The difference between the two is that the latter has classes that are overall not homogeneous. We highlight that by using the ranking provided by the Shapley effects, it is possible to select the two risk factors explaining together the highest fraction of the variance subject to a desirably reduced number of classes.

Finally, we computed the TAI in Equation (4.11). However, this index is not informative for both the PQRR and the QRPC, since all values are close to 1. For instance, the minimum and the maximum values of the index for the QRPC are 0.9999854 and 0.9999936, respectively (confirming the rating

---

[3]To avoid numerical instability in the calibration of the PQRR, a different degree of the SLP among 1, 2, and 3 is used.
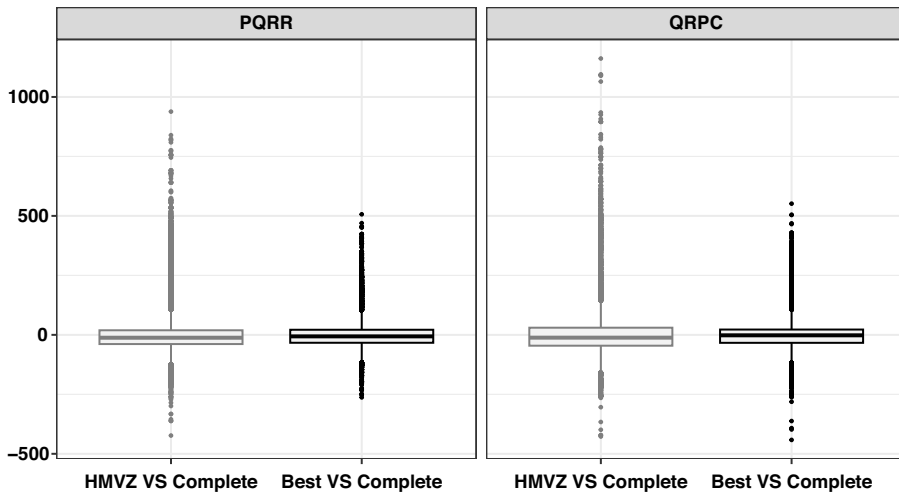
Figure 2. *Boxplots showing the differences between the premiums estimated using the complete model and two different reduced models. One of the reduced models is suggested by Heras et al. (2018), and the other is constructed using Shapley values.*

system with *vehicle body* and *agecat* as the best one in any case). From now on, we consider as the best reduced model the one including the two risk factors with the highest Shapley effects. By using the indicators in Table 3, there is evidence that the premium that a policyholder is charged in the rating class in this system is "closer" to what this policyholder would have been priced at the individual level. We provide a graphical analysis to investigate this point. In Figure 2, we compare the differences of the individual premiums found with the complete regression model (including all risk factors), with the best reduced model and the model with the two risk factors selected by Heras *et al*. (2018) (denoted with HMVZ).

Figure 2 illustrates that the model constructed with the two primary risk factors explaining the majority of the variance produces premiums that closely align with the individual premiums estimated using the complete model. This observation holds true irrespective of whether the PQRR or QRPC approach is employed for computing the individual premium. In contrast, the HMVZ model demonstrates a greater disparity from the estimated individual premiums. This is coherent with the findings from Table 3 and suggests that constructing a rating system using variables selected based on global sensitivity indices is more appropriate than relying only on *p*-values. We perform further analyses to delve deeper into the tail behavior. In Figure 3, we compute the empirical probability that the absolute value of the difference between the premiums exceeds a fixed threshold. Figure 3 shows that the tail of the best model converges to zero with a higher pace with respect to the HMVZ model. This fact confirms empirically our findings in Proposition 4.1 for a non-linear pricing model for which Chebyshev bounds cannot be obtained in a closed form. Finally, we want to demonstrate the validity of our Shapley effects approach even for the case in which the actuarial analyst wants to construct the rating system using more than two risk factors. To show this, we compute the WMCV for all possible rating systems composed of 1–6 rating factors. In Figure 4, we plot the WMCV for each rating system versus the number of selected risk factors in the system, for the insurance prices found via the PQRR approach (left panel) and QRPC approach (right panel), respectively. Let a be the total number of risk factors, $a = 6$ in our case, and let b be the number of selected risk factors, $b = 1, \ldots, 6$ in our case. The number of possible rating system we can construct for a selected number of risk factor b is given by $\binom{a}{b} = \frac{a!}{b!(a-b)!}$. For this reason, the number of gray bullets in the plot increases up to $b = 3$ and then decreases. The linked black bullets represent the rating system constructed with the risk factors found through the Shapley effects algorithm. As expected, these rating systems are the ones with lower WMCV for every number of selected risk factors. The WMCV shows a
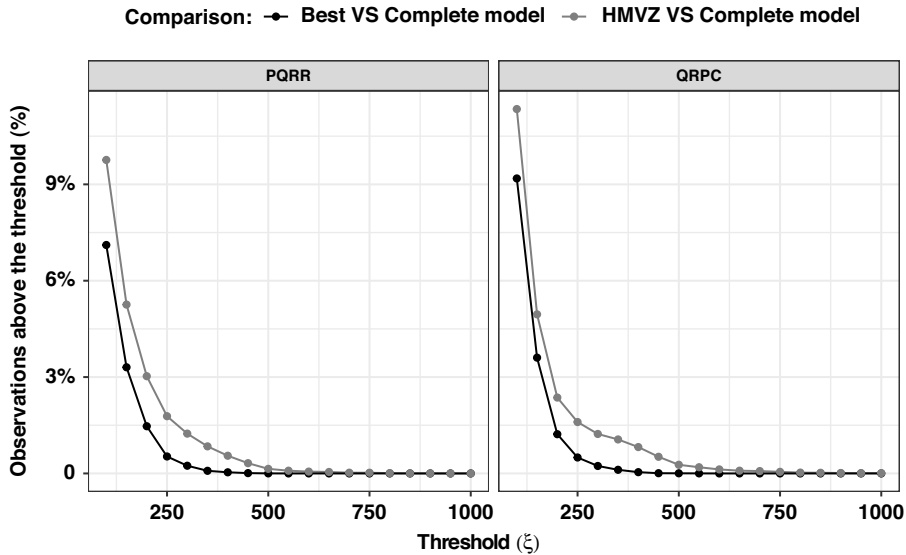
Figure 3. *Empirical probability that the individual premiums, computed with the complete model, differ from their class premiums by more than a threshold ξ, computed with the Shapley model and the HMVZ model. In both panels (PQRR on the left and QRPC on the right), given the same threshold, this probability is higher for the HMVZ model. This implies that ceteris paribus in the latter system more policyholders pay a class premium differing from the individual ones the fixed difference ξ.*

decreasing trend as the number of selected risk factors that compose the rating systems increases, since the premium computed within each risk class is clearly more accurate by using more risk factors. This is clearly intuitive. However, our approach shows that the reduction in heterogeneity varies depending on the considered risk factors. Specifically, the downward trend is more pronounced when considering up to three selected risk factors, after which the decrease becomes less significant, particularly in the case of PQRR. This is in accordance with the results obtained via the Shapley effects algorithm, since the three risk factors *agecat*, *vehicle body*, and *area* explain approximately 94% of the variance in both the PQRR model and QRPC model. Hence, considering also the fourth most important risk factors does not lead to a drastic decrease of the WMCV especially for the PQRR case. Note that if we select four risk factors, the rating system with smaller WMCV is for the PQRR case the one composed by *agecat*, *vehicle body*, *area*, and *gender*, while for the QRPC case the one composed by *agecat*, *vehicle body*, *area*, and *vehicle age*, in accordance with the ranking obtained with the Shapley effects algorithm. Finally, we remark that there is a trade-off between the number of risk classes and the overall homogeneity of the system. One could for instance directly estimate the WMCV without computing the Shapley effects. However, we suggest to estimate the Shapley effects at first and then compute the WMCV of the desired combination of the risk factors. Indeed, the WMVC needs to be estimated for all possible systems and this step has to be repeated every time the analyst wants to include more risk factors in the construction, as it happens in Figure 4.

6. Conclusions

This work presents an original procedure for the data-driven construction of rating systems using Shapley effects, with the remarkable advantage of selecting the most influential risk factors for a rating system construction.
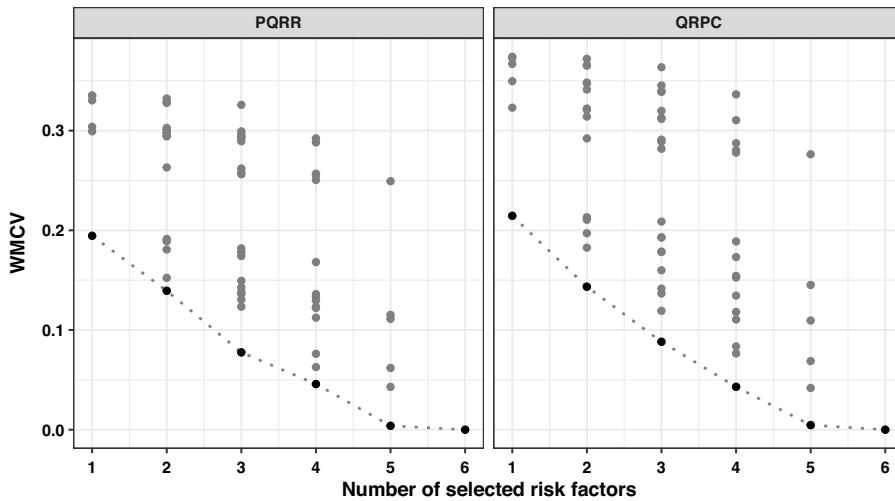
Figure 4. *WMCV for rating systems composed by a different number of risk factors. Black dots represent the WMCV of each rating system based on PQRR (left panel) and QRPC (right panel). On the x-axis, we display the number of risk factors that compose the rating systems. The dotted line links rating systems constructed using risk factors found via the Shapley effects algorithm applied to insurance prices obtained via PQRR and QRPC.*

We started with the computation of the individual quantile premium using two different methods: the QRPC approach and the PQRR approach. Nonetheless, we remark that our research does not focus on estimating individual premiums. We aim to construct rating systems based on individual premiums. As a result, our methodology is flexible and compatible with any method used to determine individual premiums. We computed the Shapley effects for these two sets of premiums. In particular, we estimated the Shapley effects by merging the cohort approach of Mase *et al*. (2019) with the Möbius inverses representation of Plischke *et al*. (2021). We consider this approach more appropriate to rank and choose the risk factors with respect to the selection based on *p*-values, since the latter does not provide information on how to rank risk factors on the basis of their importance. Moreover, *p*-values are not reliable in the case of dependent risk factors, which is typically the case in insurance datasets. Finally, in order to provide an actuarial evaluation of the quality of our approach, we introduced a new indicator, the WMCV, to compare the homogeneity of the rating systems. Results on a well-known dataset confirm the validity of our approach, since this indicator is minimized for the rating systems constructed using risk factors with the highest Shapley effects, along with other measures of homogeneity.

We believe there is broad margin for future research based on our work. To our best knowledge, the present work is the first one where a global sensitivity measure is adopted to construct a rating system. More precisely, we used Shapley effects as global importance measures. A reviewer interestingly suggested to adopt kernel methods to construct the similarity cohorts: we believe this intuition deserves a full investigation in the future. A possible drawback of the use of Shapley effects lies in their computational cost ($2^k - 1$), which is expensive for insurance datasets characterized by a large number of risk factors. Horiguchi and Pratola (2023) propose a method for estimating Shapley effects by fitting Bayesian additive regression trees. This metamodel allows the authors to obtain posterior concentration rates, which reduce the dimensionality of the risk factors, enhancing the computational feasibility of the method. Moreover, other global importance measures might be considered to construct suitable rating systems, such as those of Merz *et al.* (2021) and Makam *et al.* (2021). Another research line consists in investigating the global importance of interactions for the two pricing models, which is possible due to adoption of the Möbius inverse algorithm we have proposed. This analysis of interactions requires a dedicated future research project.

References

Baione, F. and Biancalana, D. (2019) An individual risk model for premium calculation based on quantile: A comparison between generalized linear models and quantile regression. *The North American Actuarial Journal*, **23** (4), 573–590.

Baione, F. and Biancalana, D. (2021) An application of parametric quantile regression to extend the two-stage quantile regression for ratemaking. *Scandinavian Actuarial Journal*, **2021** (2), 156–170.

Borgonovo, E., Lu, X. and Rabitti, G. (2022) Sensitivity analysis of pandemic models can support effective policy decisions. *Journal of Computational and Graphical Statistics*, 1–4. doi: 10.1080/10618600.2022.2126483. URL https://doi.org/10.1080/10618600.2022.2126483.

Borgonovo, E. and Plischke, E. (2016) Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, **248** (3), 869–887.

Da Veiga, S., Gamboa, F., Iooss, B. and Prieur, C. (2021) *Basics and Trends in Sensitivity Analysis*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

de Jong, P. and Heller, G. (2008) *Generalized Linear Models for Insurance Data. International Series on Actuarial Science*. Cambridge: Cambridge University Press.

Donnelly, C. (2014) Quantifying mortality risk in small defined-benefit pension schemes. *Scandinavian Actuarial Journal*, **2014** (1), 41–57.

EIOPA. (2009) Directive 2009/138/ec of the European parliament and of the council 25 November 2009 on the taking-up and pursuit of the business of Insurance and Reinsurance - Solvency II.

Frumento, P. and Bottai, M. (2016) Parametric modeling of quantile regression coefficient functions. *Biometrics*, **72** (1), 74–84.

Godin, F., Hamel, E., Gaillardetz, P. and Hon-Man Ng, E. (2023) Risk allocation through Shapley decompositions, with applications to variable annuities. *ASTIN Bulletin: The Journal of the IAA*. doi: 10.1017/asb.2023.7.

Henckaerts, R., Antonio, K., Clijsters, M. and Verbelen, R. (2018) A data driven binning strategy for the construction of insurance tariff classes. *Scandinavian Actuarial Journal*, **2018** (8), 681–705.

Heras, A., Moreno, I. and Vilar-Zanón, J.L. (2018) An application of two-stage quantile regression to insurance ratemaking. *Scandinavian Actuarial Journal*, **2018** (9), 753–769.

Horiguchi, A. and Pratola, M.T. (2023) Estimating Shapley effects for moderate-to-large input dimensions. arXiv: Learning. doi: arXiv:2304.03809v1.

Kaye, P. (2005) Risk measurement in insurance: A guide to risk measurement, capital allocation, and related decision support issues. Casualty Actuarial Society Discussion Paper. URL https://www.casact.org/sites/default/files/database/dpp_dpp05_05dpp1.pdf.

Koenker, R. and Bassett, G. (1978) Regression quantiles. *Econometrica*, **46** (1), 33–50.

Kudryavtsev, A.A. (2009) Using quantile regression for rate-making. *Insurance Mathematics & Economics*, **45** (2), 296–304.

Lemaire, J. (1984) An application of game theory: Cost allocation. *ASTIN Bulletin: The Journal of the IAA*, **14** (1), 61–81.

Lemaire, J. (1991) Cooperative game theory and its insurance applications. *ASTIN Bulletin: The Journal of the IAA*, **21** (1), 17–40.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N. and Lee, S.-I. (2020) From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, **2** (1), 56–67.

Mahmoudvand, R. and Oliveira, T. (2018) On the application of sample coefficient of variation for managing loan portfolio risks. In *Recent Studies on Risk Analysis and Statistical Modeling*, pp. 87–97. Springer.

Makam, V.D., Millossovich, P. and Tsanakas, A. (2021) Sensitivity analysis with χ^2-divergences. *Insurance: Mathematics and Economics*, **100**, 372–383.

Mase, M., Owen, A. B. and Seiler, B. B. (2019) Explaining black box decisions by Shapley cohort refinement. arXiv: Learning. doi: arXiv:1911.00467v2.

Mayer, M., Meier, D. and Wuthrich, M.V. (2023) Shap for actuaries: Explain any model. SSRN. doi: 10.2139/ssrn.4389797.

Merz, M., Richman, R., Tsanakas, A. and Wüthrich, M.V. (2021) Interpreting deep learning models with marginal attribution by conditioning on quantiles. SSRN. doi: 10.2139/ssrn.3809674.

Olivieri, A. and Pitacco, E. (2015) *Introduction to Insurance Mathematics: Technical and Financial Features of Risk Transfers*. Cham: Springer.

Owen, A.B. (2014) Sobol' indices and Shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, **2** (1), 245–251.

Owen, A.B. and Prieur, C. (2017) On Shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, **51** (1), 986–1002.

Plischke, E., Rabitti, G. and Borgonovo, E. (2021) Computing Shapley effects for sensitivity analysis. *SIAM/ASA Journal on Uncertainty Quantification*, **9** (4), 1411–1437.

Rabitti, G. and Borgonovo, E. (2020) Is mortality or interest rate the most important risk in annuity models? A comparison of sensitivity analysis methods. *Insurance Mathematics & Economics*, **95**, 48–58.

Saltelli, A., Marco, R., Terry, A., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S. (2008) *Global Sensitivity Analysis: The Primer*. Chichester: Wiley.

Shapley, L.S. (1953) A value for n-person games. In *Contributions to the Theory of Games II* (eds. H. Kuhn and A. Tucker), pp. 307–318. Princeton University Press.

Sobol', I. (1993) Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, **4**, 407–414.

Song, E., Nelson, B.L. and Staum, J. (2016) Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, **4** (1), 1060–1083.

Winter, E. (2002) The Shapley value. In *Handbook of Game Theory with Economic Applications* (eds. R. Aumann and S. Hart), 1 edition, vol. 3, chapter 53, pp. 2025–2054. Elsevier.