

CONVEX AND LORENZ ORDERS UNDER BALANCE CORRECTION IN NONLIFE INSURANCE PRICING: REVIEW AND NEW DEVELOPMENTS

Michel Denuit, Julien Trufin

REPRINT | 2024 / 20

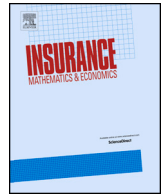
ISBA

Voie du Roman Pays 20 - L1.04.01

B-1348 Louvain-la-Neuve

Email : lidam-library@uclouvain.be

<https://uclouvain.be/en/research-institutes/lidam/isba/publication.html>



Convex and Lorenz orders under balance correction in nonlife insurance pricing: Review and new developments

Michel Denuit^a, Julien Trufin^{b,*}

^a Institute of Statistics, Biostatistics and Actuarial Science - ISBA, Louvain Institute of Data Analysis and Modeling - LIDAM, UCLouvain, Louvain-la-Neuve, Belgium

^b Department of Mathematics, Université Libre de Bruxelles (ULB), Brussels, Belgium

ARTICLE INFO

Keywords:

Tweedie deviance
Bregman divergence
Financial equilibrium
Concentration curve
Convex order
Lorenz order

ABSTRACT

By exploiting massive amounts of data, machine learning techniques provide actuaries with predictors exhibiting high correlation with claim frequencies and severities. However, these predictors generally fail to achieve financial equilibrium and thus do not qualify as pure premiums. Autocalibration effectively addresses this issue since it ensures that every group of policyholders paying the same premium is on average self-financing. Balance correction has been proposed as a way to make any candidate premium autocalibrated with the added advantage that it improves out-of-sample Bregman divergence and hence predictive Tweedie deviance. This paper proves that balance correction is also beneficial in terms of concentration curves and derives conditions ensuring that the initial predictor and its balance-corrected version are ordered in Lorenz order. Finally, criteria are proposed to rank the balance-corrected versions of two competing predictors in the convex order.

1. Introduction and motivation

Numerical evidence suggests that candidate premiums produced by machine learning tools, like boosted regression trees or neural networks, can depart a lot from observed claim totals. See e.g. Denuit et al. (2019) and the references therein. Autocalibration has been proposed as a remedy by Denuit et al. (2021), Denuit and Trufin (2021, 2023), Ciatto et al. (2023), Lindholm et al. (2023), Wüthrich (2023) and Wüthrich and Ziegel (2023). The reason is as follows: if premiums are autocalibrated then the amounts collected from policyholders paying the same premium match on average the corresponding claim totals. Subsidizing transfers are thus avoided, helping to prevent adverse selection effects. Balance correction has been proposed by Denuit et al. (2021) as a practical way to make a given predictor autocalibrated. This simple procedure consists in replacing the initial predictor with the conditional expectation of the response given the predictor.

Balance correction also decreases Bregman divergence and hence predictive Tweedie deviance. This is clearly visible in the numerical results obtained by Ciatto et al. (2023) where deviance was always reduced when switching to autocalibrated predictors with the help of the balance-correction method on the data sets under consideration. This also shows that Proposition 2.2 in Lindholm et al. (2023) extends from quadratic loss to the whole class of Bregman divergence, including

Tweedie deviance. These results demonstrate that balance correction is not only successful in restoring financial equilibrium but also improves the performances of the predictor under consideration.

In this paper, we show that balance correction is also beneficial in terms of concentration curves. Moreover, we derive criteria ensuring that a predictor compares with its balance-corrected version in the Lorenz order, and that the balance-corrected versions of two predictors are ordered in the convex order. These results demonstrate the benefits of balance correction, making it an appealing post-processing technique for insurance pricing studies.

The developments in this paper apply in any setting where balance is desirable, that is, where it is important that the sum of estimates do not deviate too much from the sum of actual observations. This naturally translates into a global balance constraint: considering that the total claim figures are representative of next-year's experience, it is important that the sum of fitted premiums matches the sum of responses taken as proxy for the total premium income, as closely as possible. In addition to global balance, balance is equally important locally in order to limit premium transfers between different categories of policyholders and to prevent adverse selection. Autocalibration appears to be the appropriate tool to translate financial equilibrium into the price list.

The remainder of this paper is organized as follows. Section 2 recalls the setting of insurance pricing and introduces the notation used

* Corresponding author.

E-mail address: julien.trufin@ulb.ac.be (J. Trufin).

throughout the paper. Eligible loss functions are also discussed there. Section 3 recalls the notion of autocalibration. Section 4 first revisits the financial equilibrium concept, central to insurance pricing, and the way it translates into balance properties at different granularity levels. Autocalibration then naturally fits this approach and appears to be an important requirement for any candidate premium. The balance correction approach for making a given predictor autocalibrated is presented. The merits of this approach are discussed. Sections 5 and 6 derive criteria to compare a predictor and its balance-corrected version in the Lorenz order and two balance-corrected predictors in the convex order.

2. Nonlife insurance pricing

2.1. Pure premiums

Given a response Y and a set of features $X_1, \dots, X_p$ gathered in the vector X , actuarial pricing aims to evaluate the pure premium as accurately as possible. This means that the target is the conditional expectation $\mu(X) = E[Y|X]$ of the response Y (claim number or claim amount) given the available information X . Henceforth, $\mu(X)$ is referred to as the true (pure) premium. Notice that in some applications, $\mu(X)$ only refers to one component of the pure premium. For instance, working in the frequency-severity decomposition of insurance losses, $\mu(X)$ can be either the expected number of insured events or the expected claim size, or severity.

The function $x \mapsto \mu(x) = E[Y|X = x]$ is unknown to the actuary, and may exhibit a complex behavior in x . This is why this function is approximated by a (working, or actual) premium $x \mapsto \pi(x)$ with a relatively simple structure compared to the unknown regression function $x \mapsto \mu(x)$. Once fitted on the training data set using an appropriate learning procedure, this produces estimates, or fitted values $\hat{\pi}(x)$ for $\mu(x)$.

The developments in this paper hold the training set as fixed. Precisely, $\hat{\pi}$ is assumed to have been obtained by minimizing the average empirical loss given some training set and all formulas in the text are meant given this training set. The respective performances of competing models can then be assessed on the basis of a validation set $\{(Y_i, X_i), i = 1, 2, \dots, n\}$, that has not been used to obtain $\hat{\pi}$. Provided the validation set comprises a large number of independent and identically distributed random vectors (Y_i, X_i) , this allows us to perform calculations with a generic random vector (Y, X) , independent of, and distributed as the observations contained in the training set.

Henceforth, we consider $Y \geq 0$. We assume that $\hat{\pi}(X)$ as well as the conditional expectation $\mu(X)$ are continuous random variables admitting positive probability density functions over $(0, \infty)$. This is generally the case when there is at least one continuous feature contained in the available information X and the function $\hat{\pi}$ is a continuously increasing function of a real score built from X . We denote as

$$F_{\hat{\pi}}(t) = P[\hat{\pi}(X) \leq t], \quad t \geq 0,$$

the distribution function of $\hat{\pi}(X)$, as $f_{\hat{\pi}}$ the corresponding probability density function, and as $F_{\hat{\pi}}^{-1}$ the associated quantile function defined as the inverse of $F_{\hat{\pi}}$.

2.2. Strictly consistent loss functions for the mean

For a convex function $\ell(\cdot)$, define the error measure for the mean as

$$L(y, m) = \ell(y) - \ell(m) - \ell'(m)(y - m). \quad (2.1)$$

Functions $L(\cdot, \cdot)$ in (2.1) are called Bregman loss functions and turn out to be particularly important in insurance ratemaking. The reason is that Savage (1971) showed, subject to weak regularity conditions, that for any loss function (2.1) and response Y , $E[L(Y, E[Y])] \leq E[L(Y, m)]$ for any m in the support of Y . The class of Bregman loss functions is then said to be strictly consistent for the (conditional) mean functional. We

refer the interested reader to Gneiting (2011) for a general overview of consistency in statistics.

Since pure premiums correspond to conditional expectations, they can thus be consistently estimated only if the expected loss is minimum for the mean response. This explains why every Bregman loss function would be eligible to produce $\hat{\pi}$ in actuarial application targeting the pure premium. See e.g. Table 1 in Krüger and Ziegel (2021) as well as Patton (2020) for examples of Bregman loss functions. In particular, Tweedie loss functions are Bregman loss functions. Performances of $\hat{\pi}$ can therefore be assessed with the help of the predictive Tweedie deviance given by

$$D(\xi, \hat{\pi}) = \begin{cases} E[\hat{\pi}(X) - Y \ln \hat{\pi}(X)] & \text{for } \xi = 1 \\ E\left[\ln \hat{\pi}(X) + \frac{Y}{\hat{\pi}(X)}\right] & \text{for } \xi = 2 \\ E\left[\frac{\hat{\pi}(X)^{2-\xi}}{2-\xi} - \frac{Y\hat{\pi}(X)^{1-\xi}}{1-\xi}\right] & \text{for } \xi > 1 \text{ and } \xi \neq 2 \end{cases}. \quad (2.2)$$

Henceforth, a predictor $\hat{\pi}_2$ outperforms $\hat{\pi}_1$ for some loss function L of the form given in (2.1) if $E[L(Y, \hat{\pi}_2(X))] \leq E[L(Y, \hat{\pi}_1(X))]$. Restricting (2.1) to Tweedie loss functions, $\hat{\pi}_2$ outperforms $\hat{\pi}_1$ if the inequality $D(\xi, \hat{\pi}_2) \leq D(\xi, \hat{\pi}_1)$ holds for some power parameter $\xi \geq 1$.

3. Autocalibration

3.1. Definition

Recall that a predictor $\hat{\pi}$ is said to be autocalibrated if $\hat{\pi}(X) = E[Y|\hat{\pi}(X)]$. Autocalibration thus ensures that the average response in a neighborhood defined with the help of the autocalibrated predictor under consideration matches the latter. See, e.g., Krüger and Ziegel (2021). Autocalibration is thus an important property in insurance pricing because it ensures that every group of policyholders paying the same premium is on average self-financing: the premium paid exactly covers the expected claims of that group.

3.2. Properties

3.2.1. Convex order and overfit

Clearly, the more an autocalibrated predictor $\hat{\pi}(X)$ varies, the more information it contains about the true premium. Thus, comparing the underlying variability appears to be important in the problem under study. The convex order is an effective probabilistic tool to assess the variability of random variables, beyond simple indicators such as standard deviations. Recall from Denuit et al. (2005, Section 3.4) that given two random variables Z and T , T is said to be smaller than Z in the convex order (denoted as $T \leq_{CX} Z$) if

$$E[g(T)] \leq E[g(Z)] \text{ for all convex functions } g, \quad (3.1)$$

provided the expectations exist.

Autocalibrated predictors are such that

$$\hat{\pi}(X) \leq_{CX} \mu(X) \leq_{CX} Y.$$

The first stochastic inequality has been obtained in Lemma 3.1 from Wüthrich (2023). The second one is well known and directly follows from Jensen inequality for conditional expectations. Thus, autocalibration implies that the predictor is less variable than the true premium $\mu(X)$, and a fortiori than the response Y , in the sense of the convex order. This important property means that premiums charged by the insurance company always induce softer relativities compared to the true ones. Stated otherwise, autocalibrated predictors never overfit.

3.2.2. Bregman dominance

Bregman dominance is defined as dominance for every Bregman loss function (2.1). Precisely, $\hat{\pi}_2$ outperforms $\hat{\pi}_1$ in terms of Bregman dom-

inance if the inequality $E[L(Y, \hat{\pi}_2(X))] \leq E[L(Y, \hat{\pi}_1(X))]$ holds true for every loss function L of the form given in (2.1). Bregman dominance is thus a particular stochastic order relation used to compare the performances of two estimators $\hat{\pi}_1$ and $\hat{\pi}_2$ for the conditional means. We refer the interested reader to Krüger and Ziegel (2021) and the references therein for an extensive presentation of this concept.

Krüger and Ziegel (2021) established in their Theorem 2.1 that $\hat{\pi}_2$ outperforms $\hat{\pi}_1$ in terms of Bregman dominance if, and only if, the inequality

$$\begin{aligned} E[(\hat{\pi}_2(X) - t)_+] + E[(Y - \hat{\pi}_2(X))I[\hat{\pi}_2(X) > t]] \\ \geq E[(\hat{\pi}_1(X) - t)_+] + E[(Y - \hat{\pi}_1(X))I[\hat{\pi}_1(X) > t]] \end{aligned} \quad (3.2)$$

holds for all t , where $I[\cdot]$ is the indicator function (equal to 1 if the event appearing within the brackets is realized and to 0 otherwise). Both sides of the latter inequality involve stop-loss premiums $E[(\hat{\pi}_k(X) - t)_+]$ of the predictors $\hat{\pi}_k$ under consideration as well as expected deficits $Y - \hat{\pi}_k(X)$ on the contracts with premium exceeding the threshold t .

Comparing predictors with the help of convex order has been proposed by Denuit et al. (2019) in their Definition 4.1. It turns out that Bregman dominance reduces to the convex order for autocalibrated predictors, as pointed out by Krüger and Ziegel (2021) in their Theorem 3.1. This is easily deduced from the following useful result.

Lemma 3.1. *For any autocalibrated predictor $\hat{\pi}$ and measurable function $g : \mathbb{R}^+ \rightarrow \mathbb{R}$, we have*

$$E[Yg(\hat{\pi}(X))] = E[\hat{\pi}(X)g(\hat{\pi}(X))].$$

Proof. This result follows from the tower property of conditional expectation, which gives

$$\begin{aligned} E[Yg(\hat{\pi}(X))] &= E[E[Yg(\hat{\pi}(X)) | \hat{\pi}(X)]] \\ &= E[E[Y | \hat{\pi}(X)]g(\hat{\pi}(X))] \\ &= E[\hat{\pi}(X)g(\hat{\pi}(X))], \end{aligned}$$

where the last step follows from autocalibration. This ends the proof. $\square$

Corollary 3.2 (Krüger and Ziegel (2021)). *Let $\hat{\pi}_1$ and $\hat{\pi}_2$ be two autocalibrated predictors. The inequality $E[L(Y, \hat{\pi}_2(X))] \leq E[L(Y, \hat{\pi}_1(X))]$ holds true for every Bregman loss function L in (2.1) if, and only if, $\hat{\pi}_1(X) \leq_{CX} \hat{\pi}_2(X)$.*

Proof. Recall that $\leq_{CX}$ can be characterized with the help of stop-loss premiums. Precisely, given two random variables Z and T ,

$$T \leq_{CX} Z \Leftrightarrow E[T] = E[Z] \text{ and } E[(T - d)_+] \leq E[(Z - d)_+] \text{ for all } d. \quad (3.3)$$

Applying Lemma 3.1 to (3.2), we get

$$E[YI[\hat{\pi}(X) > t]] = E[\hat{\pi}(X)I[\hat{\pi}(X) > t]]$$

so that

$$E[(Y - \hat{\pi}(X))I[\hat{\pi}(X) > t]] = 0$$

for every autocalibrated predictor. Only the stop-loss premiums remain in (3.2) so that we recover the stated result from (3.3). $\square$

3.2.3. Lorenz and concentration curves

The concentration curve of the response Y with respect to the predictor $\hat{\pi}$ based on the information contained in the vector X is defined as

$$CC[Y, \hat{\pi}(X); \alpha] = \frac{E[YI[\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)]]}{E[Y]}$$

for a probability level α . It turns out to be useful to assess the performances of $\hat{\pi}$, as demonstrated in Denuit et al. (2019). We refer the interested reader to Yitzhaki and Schechtman (2013) for a thorough description of concentration curves. Up to sign switches, concentration curves are referred to as cumulative accuracy profiles in binary classification, as pointed out by Wüthrich (2023). The Lorenz curve associated to the predictor $\hat{\pi}$ is defined by

$$\begin{aligned} LC[\hat{\pi}(X); \alpha] &= \frac{1}{E[\hat{\pi}(X)]} \int_0^\alpha F_{\hat{\pi}}^{-1}(u) du \\ &= \frac{E[\hat{\pi}(X)I[\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)]]}{E[\hat{\pi}(X)]} \end{aligned}$$

for a given probability level α . Now, Lemma 3.1 gives the following result.

Corollary 3.3.

$$\begin{aligned} \hat{\pi} \text{ autocalibrated} \Rightarrow CC[Y, \hat{\pi}(X); \alpha] &= \frac{E[\hat{\pi}(X)I[\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)]]}{E[Y]} \\ &= LC[Y, \hat{\pi}(X); \alpha] \end{aligned}$$

Hence, concentration and Lorenz curves coincide for autocalibrated predictors. This result established by Denuit and Trufin (2021) is recovered here in an elementary way thanks to Lemma 3.1.

4. Balance correction

4.1. Financial equilibrium and balance properties

Financial equilibrium is an important property in insurance pricing. Since risk classification is meant to vary the amounts of premium charged to policyholders, with little to no effect on reported claims, it means that the insurer must collect the same amount of premiums for the entire portfolio. Furthermore, to prevent adverse selection effects, the same property must hold locally, for meaningful groups of policyholders.

Financial equilibrium must therefore be imposed at different levels:

global balance ensures that individual fluctuations $Y - \hat{\pi}(X)$ cancel out in a sufficiently large portfolio comprising independent risks. This translates into the global balance equation

$$E[Y - \hat{\pi}(X)] = 0.$$

Considering the deficit $Y - \hat{\pi}(X)$ on the contract, this guarantees financial equilibrium of insurer's operations. Stated alternatively, this means that pricing errors $\mu(X) - \hat{\pi}(X)$ also cancel out, that is,

$$E[\mu(X) - \hat{\pi}(X)] = 0.$$

local balance prevents premium transfers from one group of policyholders to another. This ensures that premiums are competitive and avoids subsidizing transfers, helping to limit adverse selection effects. Depending on the groups over which local balance is required, this leads to different notions:

marginal balance corresponds to the method of marginal totals as well as to likelihood equations in Generalized Linear Models (GLMs) with canonical link functions and binary features X_j . If an intercept is included in the GLM score then global balance is automatically fulfilled. Furthermore, local balance also holds among policyholders with the same value for X_j . This directly results from maximum likelihood equations.

autocalibration It would be natural to impose that every group of policyholders paying the same premium is on average self-financing. This would ensure that there is no systematic cross-

financing between groups paying different amounts of premium, preventing any subsidizing solidarity and adverse selection effects. Translated into mathematical terms, this requirement corresponds to autocalibration, i.e.

$$E[Y - \hat{\pi}(X)|\hat{\pi}(X)] = 0.$$

Stated alternatively, this means that pricing errors $\mu(X) - \hat{\pi}(X)$ also cancel out, that is,

$$E[\mu(X) - \hat{\pi}(X)|\hat{\pi}(X)] = 0.$$

This shows that balance properties are central to insurance pricing and that autocalibration can be seen as the natural counterpart to marginal balance for modern machine learning ratemaking tools.

Since candidate premiums are not necessarily autocalibrated, it is interesting to provide actuaries with simple and effective ways to restore autocalibration. This is the topic of the next section.

4.2. Balance correction as a way to restore financial equilibrium

Predictors obtained by modern machine learning tools such as boosted regression trees or neural networks are generally not autocalibrated. It is therefore useful to have a practical way to make a given predictor autocalibrated. An effective method has been proposed by Denuit et al. (2021), referred to as balance correction.

Definition 4.1. The balance-corrected version $\hat{\pi}_{bc}$ of the candidate premiums $\hat{\pi}$ is defined as

$$\hat{\pi}_{bc}(X) = E[Y|\hat{\pi}(X)].$$

The predictor $\hat{\pi}_{bc}$ obtained by balance correction is always autocalibrated as shown in Proposition 3.3 in Wüthrich (2023).

Let us present two examples of balance-correction techniques.

Example 4.2 (Lift predictor). Lift charts are often used in practice to assess the performances of a candidate premium. To draw such graphs, the data set is sorted based on the values of $\hat{\pi}(X)$. The data are then bucketed into equally populated classes based on quantiles. Within each bucket, the average expected response is calculated with the help of $\hat{\pi}(X)$ as well as the actual loss Y . The average predictions and average costs are then graphed for each class. Lift charts allow the analyst to check whether the actual response monotonically increases as we move to higher buckets. Property 3.5 in Denuit et al. (2019) identifies the condition required to observe an increasing trend in such a lift chart (that is, positive regression dependence).

Let $a_0 = 0 < a_1 < a_2 < \dots$ and partition $(0, \infty)$ into intervals $I_j = (a_{j-1}, a_j)$. Define

$$\hat{\pi}_{\text{lift}}(x) = \sum_{j \geq 1} E[Y|\hat{\pi}(x) \in I_j] I[\hat{\pi}(x) \in I_j].$$

If the conditional expectations $E[Y|\hat{\pi}(x) \in I_j]$ are increasing in j , which appears to be a reasonable requirement for any candidate premium $\hat{\pi}$, we then have

$$E[Y|\hat{\pi}_{\text{lift}}(X)] = \hat{\pi}_{\text{lift}}(X)$$

so that $\hat{\pi}_{\text{lift}}$ is autocalibrated. Notice that this autocalibrated predictor is essentially the same as the one given in Definition 2.1 of Lindholm et al. (2023).

Example 4.2 considers a discrete predictor, departing from the continuity assumption made throughout this paper. Despite assuming a finite number of values, the construction relates to lift diagrams and justifies the approach proposed in the next example, leading to continuous predictors.

Example 4.3 (Moving average predictor). A locally-constant regression, or moving average approach allows the analyst to implement autocalibration, that is, to switch from $\hat{\pi}$ to $\hat{\pi}_{bc}$. See e.g. Loader (1999) for a detailed account. The difference compared to Example 4.2 is that the disjoint intervals I_j are replaced with a moving window over which responses are averaged. In fact, the score of the model is enough to compute $\hat{\pi}_{bc}$ so that the learning model is essentially used to assess proximity among individuals, before performing local averaging.

Following Denuit et al. (2021), the balance-corrected $\hat{\pi}_{bc}$ can be obtained by local regression in the appropriate GLM, recognizing the nature of the response Y : local Poisson GLM for claim counts, local Binomial GLM for propensities, local Gamma GLM for average claim sizes and local Tweedie GLM for claim totals. The choice of the response distribution is thus dictated by the format of the response. The canonical link function is adopted so that maximum-likelihood estimates replicate observed totals. An intercept-only GLM is fitted locally, on a set of observations close to the one of interest with proximity assessed with the help of $\hat{\pi}$. Precisely, $\hat{\pi}_{bc}(x)$ is obtained by averaging a percentage of the observed responses y_i whose predictions $\hat{\pi}(x_i)$ are the closest to $\hat{\pi}(x)$. This percentage is obtained by likelihood cross-validation in the appropriate GLM.

Remark 4.4 (Isotonic regression). Balance correction is not the only way to restore autocalibration. Let us mention isotonic recalibration suggested by Wüthrich and Ziegel (2023). Compared to the method described in Example 4.3, isotonic regression does not involve any hyperparameters (such as the size of the neighborhood to perform local averaging in Example 4.3) as it solves a constraint regression problem (ensuring monotonicity). Moreover, it can be shown that it does not depend on the specific choice of the loss function within the Bregman family in the sense that the same result is obtained for every loss function (2.1). This echoes locally constant, or intercept-only GLM procedures under canonical link functions in Example 4.3 which all result in moving averages. Isotonic recalibration consists in minimizing the weighted sum of squared difference between responses and the autocalibrated version of the initial predictor $\hat{\pi}$ respecting the ranking induced by $\hat{\pi}$. The solution can be given explicitly as a min-max weighted average formula.

4.3. Impact of balance correction on Bregman divergence

The numerical illustrations in Ciatto et al. (2023) suggest that predictive deviance may be smaller for $\hat{\pi}_{bc}$ compared to $\hat{\pi}$. Moreover, Proposition 2.2 from Lindholm et al. (2023) proves that the quadratic loss function is smaller for $\hat{\pi}_{bc}$ compared to $\hat{\pi}$. These are just two manifestations of the following general result: since any Bregman loss function is a strictly consistent loss function for the mean functional, balance correction necessarily improves Bregman divergence. Formally, for any loss function (2.1) we can write

$$\begin{aligned} E[L(Y, \hat{\pi}(X))] &= E[E[L(Y, \hat{\pi}(X))|\hat{\pi}(X)]] \\ &\geq E[E[L(Y, E[Y|\hat{\pi}(X)])|\hat{\pi}(X)]] \\ &= E[L(Y, E[Y|\hat{\pi}(X)])] = E[L(Y, \hat{\pi}_{bc}(X))]. \end{aligned}$$

This result is stated next, for further reference.

Proposition 4.5. We have

$$\begin{aligned} E[L(Y, \hat{\pi}(X))] &\geq E[L(Y, \hat{\pi}_{bc}(X))] \\ &\geq E[L(Y, \mu(X))] \text{ for all loss functions (2.1)} \end{aligned}$$

as well as

$$D(\xi, \mu) \leq D(\xi, \hat{\pi}_{bc}) \leq D(\xi, \hat{\pi}) \text{ for every power parameter } \xi \geq 1.$$

4.4. Impact of balance correction on concentration curves

Balance correction is also beneficial in terms of concentration curves, as shown in Proposition 4.7. To prove Proposition 4.7, we need the next lemma.

Lemma 4.6. *For any predictor $\hat{\pi}$ and measurable function $g : \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}$, we have*

$$\mathbb{E}[Y g(\hat{\pi}(X), \hat{\pi}_{bc}(X))] = \mathbb{E}[\hat{\pi}_{bc}(X) g(\hat{\pi}(X), \hat{\pi}_{bc}(X))].$$

Proof. The tower property of the conditional expectation gives

$$\begin{aligned} \mathbb{E}[Y g(\hat{\pi}(X), \hat{\pi}_{bc}(X))] &= \mathbb{E}[\mathbb{E}[Y g(\hat{\pi}(X), \hat{\pi}_{bc}(X)) | \hat{\pi}(X)]] \\ &= \mathbb{E}[\mathbb{E}[Y | \hat{\pi}(X)] g(\hat{\pi}(X), \hat{\pi}_{bc}(X))] \\ &= \mathbb{E}[\hat{\pi}_{bc}(X) g(\hat{\pi}(X), \hat{\pi}_{bc}(X))], \end{aligned}$$

which ends the proof. $\square$

Proposition 4.7. *We have $\text{CC}[Y, \hat{\pi}_{bc}(X); \alpha] = \text{LC}[\hat{\pi}_{bc}(X); \alpha] \leq \text{CC}[Y, \hat{\pi}(X); \alpha]$ for any probability level α .*

Proof. According to Property 2.4.7 in Denuit et al. (2005), we know that $\mathbb{E}[Z|A] \geq \mathbb{E}[Z|Z \leq z]$ holds for any continuous random variable Z and event A such that $\mathbb{P}[Z \leq z] = \mathbb{P}[A]$. We then also have $\mathbb{E}[Z\mathbb{I}[A]] \geq \mathbb{E}[Z\mathbb{I}[Z \leq z]]$. Applying this result with $Z = \hat{\pi}_{bc}(X)$, $\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)$ for A and $z = F_{\hat{\pi}_{bc}}^{-1}(\alpha)$, we get

$$\mathbb{E}[\hat{\pi}_{bc}(X)\mathbb{I}[\hat{\pi}_{bc}(X) \leq F_{\hat{\pi}_{bc}}^{-1}(\alpha)]] \leq \mathbb{E}[\hat{\pi}_{bc}(X)\mathbb{I}[\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)]].$$

Lemma 4.6 allows us to replace $\hat{\pi}_{bc}(X)$ with Y in the latter inequality, so that we end up with the announced result. $\square$

5. Comparing a predictor with its balance-corrected version

Balance correction is desirable in insurance pricing because it implements financial equilibrium, as explained before. In this section, we examine whether balance correction also impacts on premium differentials compared to the initial predictor. Since expected values of $\hat{\pi}(X)$ and $\hat{\pi}_{bc}(X)$ generally differ, we cannot compare them with the help of $\leq_{\text{CX}}$ but $\frac{\mathbb{E}[Y]}{\mathbb{E}[\hat{\pi}(X)]} \hat{\pi}(X)$ and $\hat{\pi}_{bc}(X)$ could possibly be ranked in the convex order. Here, the factor $\frac{\mathbb{E}[Y]}{\mathbb{E}[\hat{\pi}(X)]}$ restores global balance. This corresponds to the Lorenz order $\leq_{\text{Lorenz}}$. Formally,

$$\begin{aligned} \hat{\pi}(X) \leq_{\text{Lorenz}} \hat{\pi}_{bc}(X) &\Leftrightarrow \frac{\mathbb{E}[Y]}{\mathbb{E}[\hat{\pi}(X)]} \hat{\pi}(X) \leq_{\text{CX}} \hat{\pi}_{bc}(X) \Leftrightarrow \frac{\hat{\pi}(X)}{\mathbb{E}[\hat{\pi}(X)]} \\ &\leq_{\text{CX}} \frac{\hat{\pi}_{bc}(X)}{\mathbb{E}[\hat{\pi}_{bc}(X)]}. \end{aligned}$$

Lorenz order can be equivalently defined by pointwise comparison of Lorenz curves.

Balance correction consists in replacing the predictor $\hat{\pi}$ with a function of it, corresponding to the conditional expectation of the response given the predictor. The class of functions inducing Lorenz dominance is known from Theorems 4.1–4.2 in Arnold (1991).

Proposition 5.1. *Define the function $m(\cdot)$ as $m(t) = \mathbb{E}[Y | \hat{\pi}(X) = t]$. Assume that $m(\cdot)$ is continuous and strictly increasing. Then,*

$$t \mapsto \frac{m(t)}{t} \text{ non-decreasing} \Rightarrow \hat{\pi}(X) \leq_{\text{Lorenz}} \hat{\pi}_{bc}(X)$$

while

$$t \mapsto \frac{m(t)}{t} \text{ non-increasing} \Rightarrow \hat{\pi}_{bc}(X) \leq_{\text{Lorenz}} \hat{\pi}(X).$$

The sufficient conditions in Proposition 5.1 are best explained using loss ratios. Indeed,

$$\frac{m(t)}{t} = \mathbb{E}\left[\frac{Y}{\hat{\pi}(X)} \middle| \hat{\pi}(X) = t\right]$$

involves the ratio of the claim response Y to candidate premium $\hat{\pi}(X)$. They can also be interpreted as a monotonicity condition on

$$\frac{m(t)}{t} = \frac{\mathbb{E}[Y | \hat{\pi}(X) = t]}{\mathbb{E}[Y | \hat{\pi}_{bc}(X) = t]},$$

that is, on expected claims given the same value t of the predictors, before and after balance correction.

Proposition 5.1 assumes that $m(\cdot)$ is increasing. This property is referred to as regression dependence. We know from Property 3.5 in Denuit et al. (2019) that this is equivalent to the convexity of the concentration curve. This assumption may be restrictive in some cases. For instance, if the initial predictor $\hat{\pi}$ is poorly accurate then its transformation through $m(\cdot)$ may result in a change in the way premiums are ordered. The monotonicity of the function $m(\cdot)$ can be formally tested using the tools developed after Bowman et al. (1998). If the initial predictor $\hat{\pi}$ is sufficiently informative, in the sense that it is strongly correlated to the response, we may expect that $m(\cdot)$ is increasing. This is the situation considered in this paper, where we start from an initial predictor $\hat{\pi}$ exhibiting a strong positive dependence with the response but which does not qualify as a premium because it does not fulfill balance requirements. In that case, the regression dependence assumption appears to be reasonable.

If $m(\cdot)$ is continuous and strictly increasing then

$$\text{CC}[Y, \hat{\pi}_{bc}(X); \alpha] = \text{CC}[Y, \hat{\pi}(X); \alpha] \text{ for all probability levels } \alpha$$

since $\hat{\pi}_{bc}(X) \leq F_{\hat{\pi}_{bc}}^{-1}(\alpha)$ if, and only if, $\hat{\pi}(X) \leq F_{\hat{\pi}}^{-1}(\alpha)$ for any probability level α in this case. Moreover, $\hat{\pi}_{bc}$ being autocalibrated, its concentration curve is equal to its Lorenz curve. This leads to the following result.

Proposition 5.2. *If $m(\cdot)$ is continuous and strictly increasing then*

$$\hat{\pi}(X) \leq_{\text{Lorenz}} \hat{\pi}_{bc}(X) \Leftrightarrow \text{LC}[\hat{\pi}(X); \alpha]$$

$$\geq \text{CC}[Y, \hat{\pi}(X); \alpha] \text{ for all probability levels } \alpha$$

while

$$\hat{\pi}_{bc}(X) \leq_{\text{Lorenz}} \hat{\pi}(X) \Leftrightarrow \text{LC}[\hat{\pi}(X); \alpha]$$

$$\leq \text{CC}[Y, \hat{\pi}(X); \alpha] \text{ for all probability levels } \alpha.$$

In the numerical examples of Denuit et al. (2019) and in the case study of Denuit et al. (2023), it is observed that $\text{LC}[\hat{\pi}(X); \alpha] \leq \text{CC}[Y, \hat{\pi}(X); \alpha]$ for all probability levels α . In these examples, Proposition 5.2 shows that when $m(\cdot)$ is continuous and strictly increasing, we end up with $\hat{\pi}_{bc}(X) \leq_{\text{Lorenz}} \hat{\pi}(X)$, which means that balance correction softens premium differentials induced by the initial predictor. Notice that as balance correction is a way to restore autocalibration, it is also a way to remove potential overfitting since autocalibrated predictors never overfit (see Section 3.2.1).

6. Comparing balance-corrected predictors

Assume that we have two predictors $\hat{\pi}_1$ and $\hat{\pi}_2$. Both attempt to predict the unknown pure premium. These predictors may differ in their functional form and/or in the information on which they are based. Since balance corrections essentially consist in using the candidate premium to rank risk profiles from the cheapest one to the most expensive one, it is clear that if $\hat{\pi}_1$ and $\hat{\pi}_2$ are comonotonic then their ranks coincide and $\hat{\pi}_{bc,1} = \hat{\pi}_{bc,2}$.

For any reasonable candidate premium, regression dependence must hold true. The next result then identifies a situation where the balance-corrected version of two predictors is ordered in the convex sense. This

is particularly useful since Bregman dominance then also holds because they are both autocalibrated, as shown in Corollary 3.2.

Proposition 6.1. *Let $m_k(t) = E[Y | \hat{\pi}_k(X) = t]$ for $k \in \{1, 2\}$. If $m_1(\cdot)$ and $m_2(\cdot)$ are continuous and strictly increasing then*

$$CC[Y, \hat{\pi}_2(X); \alpha] \leq CC[Y, \hat{\pi}_1(X); \alpha] \text{ for all } \alpha \Rightarrow \hat{\pi}_{bc,1}(X) \leq_{CX} \hat{\pi}_{bc,2}(X).$$

Proof. By the same reasoning as for Proposition 5.2, we have

$$\begin{aligned} CC[Y, \hat{\pi}_k(X); \alpha] &= CC[Y, \hat{\pi}_{bc,k}(X); \alpha] \\ &= LC[\hat{\pi}_{bc,k}(X); \alpha] \text{ for all probability levels } \alpha, \end{aligned}$$

from which it follows that $\hat{\pi}_{bc,1}(X) \leq_{\text{Lorenz}} \hat{\pi}_{bc,2}(X)$ holds true. The announced result then follows since $E[\hat{\pi}_{bc,1}(X)] = E[\hat{\pi}_{bc,2}(X)] = E[Y]$. $\square$

Proposition 6.1 can also be seen as a consequence of Proposition 3.1 in Denuit (2010). A single sign-change property is enough to ensure that the balance-corrected versions of $\hat{\pi}_1$ and $\hat{\pi}_2$ are ordered in the convex sense. See Proposition 2.1 in Denuit (2010).

7. Discussion

This paper aims to summarize recent developments related to autocalibration and to derive new results illustrating its relevance for insurance pricing. All results are established, or re-established in case of existing ones in a simple, transparent and unified way. The paper also stresses their intimate relation to statistical learning techniques used by actuaries. In terms of novelty, three main results have been derived in this paper. First, it is established that balance-corrected premium leads to smaller concentration curves. Second, conditions are derived under which the initial predictor and its balance-corrected version are ordered in Lorenz order. Third, the paper provides actuaries with some criteria that can be used to rank two competing balance-corrected predictors.

To end with, let us briefly comment on the exposure included in the analysis. This paper adopted the traditional approach (in motor insurance, for instance) with the duration of the cover period treated as a known constant and included with the help of an offset in Poisson regression for claim counts (or equivalently, used as a weight when claim frequencies are studied). This approach is in line with the situations commonly encountered in practice, where policyholder's profile changes during the observation period, because he or she buys a new vehicle or moves to another place, for instance. The amount of premium is adapted accordingly and the cover is not interrupted. This policyholder then appears twice in the database, once with the profile before the change and a second time with the profile after the change. If any, claims are recorded with the profile at occurrence.

Lindholm et al. (2023) adopted another point of view and explicitly allowed for randomness in risk exposure. These authors study the case of random duration resulting from contract termination before the end of the agreed term. Think for instance to a change in ownership. The cancellation probability as well as the reduced contract duration may also vary according to policyholder's profile. Another situation where this approach appears to be relevant relates to home insurance, where the property value is assessed with the help of a questionnaire filled by the policyholder. To avoid the application of the proportional rule, most companies indeed provide policyholders with a property appraisal form. If the form is filled in good faith, the premium is determined on the basis of the value resulting from the answers to the various questions (such as number of bedrooms, building materials, etc.), and the company will not be able to invoke the proportional rule clause even if it discovers after the claim that the value of the property has been underestimated. If the exposure is taken to be the value of the property then it must be treated as a random variable because it only results

from a prediction made by the company on the basis of the form, and may well depart from the actual sum insured. It would be interesting to extend the results derived in this paper to the change of measure with respect to exposure advocated by Lindholm et al. (2023). We leave this topic for future research.

CRedit authorship contribution statement

Michel Denuit: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Julien Trufin:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Data availability

No data was used for the research described in the article.

Acknowledgements

The authors would like to express their gratitude to the anonymous Referees and Handling Editor whose comments have been extremely useful to revise a previous version of the present work.

References

- Arnold, B.C., 1991. Preservation and attenuation of inequality as measured by the Lorenz order. In: *Stochastic Orders and Decision Under Risk*. In: *Lecture Notes-Monograph Series*, vol. 19, pp. 25–37.
- Bowman, A.W., Jones, M.C., Gijbels, I., 1998. Testing monotonicity of regression. *J. Comput. Graph. Stat.* 7, 489–500.
- Ciatto, N., Verelst, H., Trufin, J., Denuit, M., 2023. Does autocalibration improve goodness of fit? *Eur. Actuar. J.* 13, 479–486.
- Denuit, M., 2010. Positive dependence of signals. *J. Appl. Probab.* 47, 893–897.
- Denuit, M., Charpentier, A., Trufin, J., 2021. Autocalibration and Tweedie-dominance for insurance pricing with machine learning. *Insur. Math. Econ.* 101, 485–497.
- Denuit, M., Dhaene, J., Goovaerts, M.J., Kaas, R., 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*. Wiley, New York.
- Denuit, M., Huyghe, J., Verdebout, T., Trufin, J., 2023. Testing for auto-calibration with Lorenz and concentration curves. Working Paper.
- Denuit, M., Sznajder, D., Trufin, J., 2019. Model selection based on Lorenz and concentration curves, Gini indices and convex order. *Insur. Math. Econ.* 89, 128–139.
- Denuit, M., Trufin, J., 2021. Lorenz curve, Gini coefficient, and Tweedie dominance for autocalibrated predictors. Working Paper available from dial.uclouvain.be.
- Denuit, M., Trufin, J., 2023. Model selection with Pearson's correlation, concentration and Lorenz curves under autocalibration. *Eur. Actuar. J.* 13, 871–878.
- Gneiting, T., 2011. Making and evaluating point forecasts. *J. Am. Stat. Assoc.* 106, 746–762.
- Krüger, F., Ziegel, J.F., 2021. Generic conditions for forecast dominance. *J. Bus. Econ. Stat.* 39, 972–983.
- Lindholm, M., Lindskog, F., Palmquist, J., 2023. Local bias adjustment, duration-weighted probabilities, and automatic construction of tariff cells. *Scand. Actuar. J.* 2023 (10), 946–973.
- Loader, C., 1999. *Local Regression and Likelihood*. Springer, New York.
- Patton, A.J., 2020. Comparing possibly misspecified forecasts. *J. Bus. Econ. Stat.* 38, 796–809.
- Savage, L.J., 1971. Elicitation of personal probabilities and expectations. *J. Am. Stat. Assoc.* 66, 783–810.
- Wüthrich, M.V., 2023. Model selection with Gini indices under auto-calibration. *Eur. Actuar. J.* 13, 469–477.
- Wüthrich, M.V., Ziegel, J., 2023. Isotonic recalibration under a low signal-to-noise ratio. *Scand. Actuar. J.* <https://doi.org/10.1080/03461238.2023.2246743>.
- Yitzhaki, S., Schechtman, E., 2013. *The Gini Methodology: A Primer on Statistical Methodology*. Springer.