On the Properties of the Primary Loss and the Excess Loss in NCCI's Experience Rating Plan

by Liang Hong

ABSTRACT

Split credibility has been used in practice for several decades, though its foundational theory has been investigated only recently. This paper studies the properties of the primary loss and the excess loss in the split experience plan of the National Council on Compensation Insurance (NCCI). We first revisit the claim that the excess loss is more volatile than the total loss. We show that this claim holds in the collective risk model with an arbitrary frequency distribution, generalizing an extant result in which the frequency distribution is a Poisson distribution. We also show that the primary loss is less volatile than the total loss. Next, we show that the previously established ordering of the coefficients of variation of the primary loss, the excess loss, and the total loss also holds in a more general model. Finally, we investigate the covariance and correlation coefficient between the primary loss and the excess loss. We also discuss some potential applications of our results. The paper concludes with some conjectures.

KEYWORDS

Workers' compensation, credibility theory, split point, state limit, coefficient of variation, total loss, covariance, correlation coefficient, collective risk model

1. Introduction

The American workers' compensation system is one of the most successful social programs in the country. Many U.S. states are currently using the experience rating plan of the National Council on Compensation Insurance (NCCI). In the current NCCI experience rating plan, a cap, called the *state limit*, is first applied to each individual claim. Next, a cutoff point, called the *split point*, is applied to the capped claim; that is, the capped claim is split into two components—(1) the primary loss, which reflects frequency, and (2) the excess loss, which reflects severity—and then two different credibilities are assigned to these two components. In the literature, such a practice is referred to as *split credibility*. Split credibility has been used in practice for a few decades and studied by several authors, such as Gillam (1992), Mahler (1987), and Venter (1987). Recently, Robbin (2013) initiated a rigorous study of split credibility and derived several interesting results. The purpose of this paper is to investigate the properties of the primary loss and the excess loss in a split rating plan.

Robbin (2013) showed that the coefficient of variation (CV) of the total loss is no greater than that of the excess loss in the collective risk model[1] when the frequency distribution is a Poisson distribution. We show that this conclusion holds for an arbitrary frequency distribution as well. In addition, we show that the CV of the primary loss is no greater than that of the total loss. These steps establish the ordering of the CVs of the primary loss, the excess loss, and the total loss.

Next, we point out that the usual credibility theory framework implies a model that is more general than the one in the collective risk model. We call such a model the *generalized collective risk model*. Therefore, it is natural to ask whether the established ordering of the CVs of the primary loss, the excess loss, and the total loss in the collective risk model still holds in this generalized collective risk model. We show that the answer is affirmative.

Finally, we study the covariance and correlation coefficient between the primary loss and the excess loss, establishing several interesting results. For example, (1) neither the covariance nor the correlation coefficient between the primary loss and the excess loss of a claim is a monotonic function of the split point; (2) the covariance between the primary loss of a claim and the excess loss of a different claim is a nondecreasing function of the state limit; (3) the covariance between the primary loss and the excess loss of a claim is a nondecreasing function of the state limit, but the correlation coefficient between the primary loss and the excess loss of a claim is a nonincreasing function of the state limit; (4) the primary loss and the excess loss of a single claim are nonnegatively correlated, as are the primary loss of one claim and the excess loss of another. Table 4.6 summarizes the key results along this line.

The paper concludes with two conjectures: (1) the correlation coefficient between the primary loss of one claim and the excess loss of another is a nondecreasing function of both the state limit and the split point; (2) the correlation coefficient between the primary loss and the excess loss of a claim is a nondecreasing function of the state limit.

The remainder of the paper is organized as follows. Section 2 establishes that the primary loss is less volatile than the total loss, which in turn is less volatile than the excess loss, in the collective risk model. Section 3 establishes the same ordering of the primary loss, the excess loss, and the total loss in the generalized collective risk model. Section 4 investigates the covariance and correlation between the primary loss and the excess loss. Section 5 concludes the paper with a summary and two conjectures.

---

[1]To our knowledge, the term *collective risk model* has been defined variously in the prior literature. Our definition agrees with definition 6.1 in Klugman, Panjer, and Willmot (2004) and the definition in Chapter 13 of Cunningham, Herzog, and London (2008). The definition in Heckman and Meyers (1983) carries the same name but is slightly more general and is similar to what we refer to as the *generalized collective risk model* (see definition in Section 3).

2. Coefficients of variation of the primary loss, the excess loss, and the total loss in the collective risk model

Following Robbin (2013), we will use the CV to measure the volatility of a random variable. Robbin (2013) noticed an intuitive justification for split credibility: if the total loss is split into the primary loss and the excess loss, both are less volatile than the total loss, making them more credible than the total loss. Assuming that the (claim) frequency distribution is a Poisson distribution, Robbin proved that the CV of the total loss is no greater than the CV of the excess loss in the collective risk model, thus demonstrating that the intuitive justification is incorrect. But it is still desirable to know whether this intuitive justification may be valid for some other frequency distributions. Since Robbin's result concerns only the excess loss and the total loss, it is also natural to ask whether the total loss is less volatile than the primary loss. The purpose of this section is to investigate these two problems. We establish that the CV of the total loss is never greater than the CV of the excess loss for an arbitrary frequency distribution. Then we show that the CV of the primary loss is no greater than that of the total loss.

To start, we consider the collective risk model. Let N be the number of claims and X_i be the loss amount of the ith claim ($i = 1, 2, \ldots, N$). We specify that $A = X_1 + X_2 + \ldots + X_N$; that is, A denotes the aggregate loss. We assume that all X_is are independent and identically distributed (iid) with an absolutely continuous cumulative distribution function (CDF) of $F(x)$ and a survival function of $S(x)$. We also assume that each X_i is independent of N. Let C denote the state limit and K denote the split point under a split rating plan. Without loss of generality, we assume that $0 \le K \le C$. Then for the ith claim, the primary loss and the excess loss are given by

$$X_{ip} = \begin{cases} X_i, & \text{if } X_i \le K; \\ K, & \text{if } X_i > K; \end{cases}$$

and

$$X_{ie} = \begin{cases} 0, & \text{if } X_i \le K; \\ X_i - K, & \text{if } K < X_i \le C; \\ C - K, & \text{if } X_i > C, \end{cases}$$

respectively. For aggregate loss A, the primary loss, A_p, and the excess loss, A_e, are defined by

$$A_p = X_{1p} + \ldots + X_{Np},$$

and

$$A_e = X_{1e} + \ldots + X_{Ne},$$

respectively. (To avoid complicated notation, we will often suppress the index i when we refer to an individual claim.) The mean and variance of N will be denoted by μ_N and σ_N^2, respectively; μ_X, σ_X, μ_{X_p}, σ_{X_p}, μ_{X_e}, and σ_{X_e} should be interpreted similarly. For aggregate loss A, the CVs of the primary loss, the excess loss, and the total loss will be denoted by CV_{A_p}, CV_{A_e}, and CV_A, respectively; for an individual loss X, the CVs of the primary loss, the excess loss, and the total loss will be denoted by CV_{X_p}, CV_{X_e}, and CV_X, respectively.

Theorem 2.1. *For aggregate loss A in the collective risk model, the CV of the total loss is no greater than the CV of the excess loss; that is,*

$$CV_A \le CV_{A_e}.$$

In particular, for a single claim, X, the CV of the total loss is no greater than the CV of the excess loss; that is,

$$CV_X \le CV_{X_e}.$$

Proof:

The square of the CV of aggregate loss A is given by

$$(CV_A)^2 = \frac{\mu_X^2 \sigma_N^2 + \mu_N \sigma_X^2}{(\mu_X \mu_N)^2} = \frac{\sigma_N^2}{\mu_N^2} + \frac{1}{\mu_N} \frac{\sigma_X^2}{\mu_X^2}.$$

The square of the CV of the excess loss equals

$$(CV_{A_e})^2 = \frac{\mu_{X_e}^2\sigma_N^2 + \mu_N\sigma_{X_e}^2}{(\mu_{X_e}\mu_N)^2} = \frac{\sigma_N^2}{\mu_N^2} + \frac{1}{\mu_N}\frac{\sigma_{X_e}^2}{\mu_{X_e}^2}.$$

Therefore, it suffices to show that

$$\frac{\sigma_X^2}{\mu_X^2} \le \frac{\sigma_{X_e}^2}{\mu_{X_e}^2}. \tag{2.1}$$

To this end, we write Equation (2.1) as

$$\frac{E[X^2]-\mu_X^2}{\mu_X^2} \le \frac{E[X_e^2]-\mu_{X_e}^2}{\mu_{X_e}^2},$$

which is equivalent to

$$\frac{E[X^2]}{\mu_X^2} \le \frac{E[X_e^2]}{\mu_{X_e}^2}. \tag{2.2}$$

We will establish Equation (2.2) by showing that the function $G(K) = E[X_e^2]/\mu_{X_e}^2$ is nondecreasing on the positive real line R_+. The derivatives of $E[X_e^2]$ and $\mu_{X_e}^2$ are

$$\frac{\partial}{\partial K}E[X_e^2] = \frac{\partial}{\partial K}\left[\int_K^C (x-K)^2 dF(x) + (C-K)^2 S(C)\right]$$
$$= -2\mu_{X_e},$$

and

$$\frac{\partial}{\partial K}E(\mu_{X_e}^2) = \frac{\partial}{\partial K}\left[\int_K^C (x-K)dF(x) + (C-K)S(C)\right]^2$$
$$= 2\left[\int_K^C (x-K)dF(x) + (C-K)S(C)\right]$$
$$\left[\int_K^C -1\,dF(x) - S(C)\right]$$
$$= -2\mu_{X_e}S(K),$$

respectively. It follows that

$$G'(K) = \frac{2}{\mu_{X_e}^3}\left(S(K)E[X_e^2] - \mu_{X_e}^2\right).$$

Therefore, the whole matter boils down to showing that

$$\mu_{X_e}^2 \le S(K)E[X_e^2]. \tag{2.3}$$

To establish Equation (2.3), we write μ_X^2 as

$$\mu_{X_e} = \int_K^\infty 1 \cdot g(x)dF(x),$$

where

$$g(x) = \begin{cases} 0, & \text{if } x \le K; \\ x-K, & \text{if } K < x \le C; \\ C-K, & \text{if } x > C, \end{cases}$$

and apply the Cauchy-Schwarz inequality. ■

Theorem 2.2. *For aggregate loss A in the collective risk model, the CV of the primary loss is no greater than the CV of the total loss; that is,*

$$CV_{A_p} \le CV_A.$$

In particular, for a single loss, X, the CV of the primary loss is no greater than the CV of the total loss; that is,

$$CV_{X_p} \le CV_X.$$

Proof:

Similar to the proof of Theorem 2.1, we have

$$(CV_A)^2 = \frac{\sigma_N^2}{\mu_N^2} + \frac{1}{\mu_N}\frac{\sigma_X^2}{\mu_X^2},$$

and

$$(CV_{A_p})^2 = \frac{\sigma_N^2}{\mu_N^2} + \frac{1}{\mu_N}\frac{\sigma_{X_p}^2}{\mu_{X_p}^2}.$$

Therefore, it suffices to show that

$$\frac{\sigma_{X_p}^2}{\mu_{X_p}^2} \le \frac{\sigma_X^2}{\mu_X^2},$$

which is equivalent to

$$\frac{E\left[X_p^2\right]}{\mu_{X_p}^2} \leq \frac{E\left[X^2\right]}{\mu_X^2}.$$

To this end, we specify

$$G(K) = \frac{E\left[X_p^2\right]}{\mu_{X_p}^2},$$

where K is the split point. Then $\lim_{K\to\infty} E[X_p^2]/\mu_{X_p}^2 = E[X^2]/\mu_X^2$. Thus, we need only to show that $G(K)$ is nondecreasing on R_+. We have

$$\begin{aligned}\frac{\partial}{\partial K} E\left[X_p^2\right] &= \frac{\partial}{\partial K}\left[\int_0^K x^2 dF(x) + K^2 S(K)\right] \\ &= 2KS(K),\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial K}\left(\mu_{X_p}^2\right) &= 2\mu_{X_p}\frac{\partial}{\partial K}\left[\int_0^K x dF(x) + KS(K)\right] \\ &= 2\mu_{X_p} S(K).\end{aligned}$$

It follows that

$$G'(K) = \frac{2S(K)}{\mu_{X_p}^3}\left(K\mu_{X_p} - E\left[X_p^2\right]\right) \geq 0,$$

where the last inequality follows from the fact that

$$\begin{aligned}K\mu_{X_p} &= K\left(\int_0^K x dF(x) + KS(K)\right) \\ &= K\int_0^K x dF(x) + K^2 S(K) \\ &\geq \int_0^K x^2 dF(x) + K^2 S(K) = E\left[X_p^2\right].\end{aligned}$$

■

Summarizing Theorem 2.1 and Theorem 2.2, we have the following conclusion:

Theorem 2.3. *For aggregate loss A in the collective risk model, the CVs of the primary loss, the excess loss, and the total loss satisfy the following relationship:*

$$CV_{A_p} \leq CV_A \leq CV_{A_e}.$$

In particular, for a single claim X, the CVs of the primary loss, the excess loss, and the total loss satisfy the following inequality:

$$CV_{X_p} \leq CV_X \leq CV_{X_e}.$$

Theorem 2.3 establishes the ordering of the CVs of the primary loss, the excess loss, and the total loss. It shows that the primary loss is always less volatile than the total loss. However, this does not mean the primary loss is more credible than the total because, as Robbin (2013) pointed out, volatility alone does not determine credibility. Also, Theorem 2.3 implies neither that the total loss is riskier than the primary loss nor that the excess loss is riskier than the total loss. The reason is that "risk" and "volatility" are not synonymous; see, for example, Brockett and Garven (1998). What Theorem 2.3 does tell us is that the current NCCI experience rating plan splits each aggregate loss into two components, with one being less volatile and the other more volatile. Therefore, the intuitive justification discussed in Section 2 is incorrect regardless of the distribution of the claim frequency. For the correct justification, see Robbin (2013).

In the collective risk model, all random claim-severity variables are assumed to be iid. This means that all policyholders in the pool are from one homogeneous risk class. However, this assumption is rarely met in reality; see, for example, Bühlmann and Gisler (2005) for more discussion along this line. Therefore, it makes sense to relax this iid assumption so that the uncertainty of the policyholder's risk class is taken into consideration. For this reason, one might want to investigate whether the key results in this section will still hold if the random claim-severity variables are conditionally iid. The next section is devoted to this task.

3. Coefficients of variation of the primary loss, the excess loss, and the total loss in the generalized collective risk model

In the usual setup of credibility theory, each of the observed claims, $X_1, \ldots, X_n, \ldots$, is assumed to belong to a risk class indicated by a *risk parameter*, θ, and all the risk classes are described by a risk family, Θ, called the *parameter space*. Given a risk parameter $\theta \in \Theta$, $X_1, \ldots, X_n, \ldots$ are assumed to be iid. Though $X_1, \ldots, X_n, \ldots$ are assumed to be conditionally iid, they need not be independent. Indeed, the conditional covariance formula implies that the covariance between X_i and X_j equals

$$Var_\theta[E(X_1|\theta)]$$

in such a model; see, for example, Bühlmann and Gisler (2005).[2] Therefore, we consider the following model: claim severity random variables $X_1, \ldots, X_n, \ldots$ are conditionally iid, given that the risk parameter, θ, and the random frequency variable, N, are independent of all the X_is. To distinguish this model from the collective risk model in Section 2, we call this model the *generalized collective risk model.*[3] Thus it is natural to ask whether the conclusion of Theorem 2.3 still holds in the generalized collective risk model. The rest of this section is devoted to showing that the answer is affirmative. To lighten the notation, we will use μ_X^θ and σ_X^θ to denote $E[X|\theta]$ and $Var(X|\theta)$, respectively. $\mu_{X_e}^\theta$, $\sigma_{X_e}^\theta$, $\mu_{X_p}^\theta$, and $\sigma_{X_p}^\theta$ are to be interpreted in a similar manner.

[2]Technically, $X_1, \ldots, X_n, \ldots$ are said to be *exchangeable*. That is, for every finite collection $X_{i1}, \ldots, X_{in}$, every permutation of $(X_{i1}, \ldots, X_{in})$ has the same joint distribution. This follows from the de Finetti theorem; see, for example, Theorem 1.49 of Schervish (1995). In other words, the basic framework of credibility theory entails an exchangeable model.

[3]The "collective risk model" in Heckman and Meyers (1983) puts a priority on the frequency distribution; otherwise, it would be identical to our "generalized collective risk model."

Theorem 3.1. *For aggregate loss A in the generalized collective risk model, the CV of the total loss is no greater than the CV of the excess loss; that is,*

$$CV_A \le CV_{A_e}.$$

Proof:

Let $A = X_1 + X_2 + \ldots + X_N$. Then the mean of A may be calculated using the iterated expectation formula as follows:

$$\begin{aligned}\mu_A &= E_N[E[A|N]] = E[NE[X]]\\ &= E[N]E[X] = \mu_N \mu_X.\end{aligned}$$

Also, the mean of A, given that $N = n$, equals

$$E[A|N=n] = nE[X] = n\mu_X,$$

and the variance of A, given that $N = n$, equals

$$\begin{aligned}Var[A|N=n] &= Var[X_1 + \ldots + X_n]\\ &= E_\theta[Var[X_1 + \ldots X_n|\theta]]\\ &\quad + Var_\theta[E[X_1 + \ldots X_n|\theta]]\\ &= nE_\theta[\sigma_X^\theta] + n^2 Var_\theta[\mu_X^\theta].\end{aligned}$$

It follows from the previous two displays that

$$\begin{aligned}\sigma_A^2 &= E_N[Var(A|N)] + Var_N[E(A|N)]\\ &= \sigma_N^2\mu_X^2 + \mu_N E_\theta[\sigma_X^\theta] + E[N^2]Var_\theta[\mu_X^\theta].\end{aligned}$$

Therefore,

$$\begin{aligned}(CV_A)^2 &= \frac{\sigma_N^2\mu_X^2 + \mu_N E_\theta[\sigma_X^\theta] + E[N^2]Var_\theta[\mu_X^\theta]}{\mu_N^2\mu_X^2}\\ &= \left(\frac{\sigma_N}{\mu_N}\right)^2 + \frac{\mu_N E_\theta[\sigma_X^\theta] + E[N^2]Var_\theta[\mu_X^\theta]}{\mu_N^2\mu_X^2}.\end{aligned}$$

Likewise, for the excess loss we have

$$\mu_{A_e} = \mu_N \mu_{X_e};$$

$$\sigma_{A_e}^2 = \sigma_N^2\mu_{X_e}^2 + \mu_N E_\theta[\sigma_{X_e}^\theta] + E[N^2]Var_\theta[\mu_{X_e}^\theta];$$

$$(CV_{A_e})^2 = \left(\frac{\sigma_N}{\mu_N}\right)^2 + \frac{\mu_N E_\theta[\sigma^\theta_{X_e}] + E[N^2]Var_\theta[\mu^\theta_{X_e}]}{\mu_N^2\mu_{X_e}^2}.$$

To show that $CV_A \le CV_{A_e}$, we need only to establish that

$$\frac{\mu_N E_\theta[\sigma^\theta_{X}] + E[N^2]Var_\theta[\mu^\theta_{X}]}{\mu_{X}^2} \le \frac{\mu_N E_\theta[\sigma^\theta_{X_e}] + E[N^2]Var_\theta[\mu^\theta_{X_e}]}{\mu_{X_e}^2}. \quad (3.1)$$

To this end, we write

$$G(K) = \frac{\mu_N E_\theta[\sigma^\theta_{X_e}] + E[N^2]Var_\theta[\mu^\theta_{X_e}]}{\mu_{X_e}^2}.$$

Since $G(0) = (CV_A)^2$, it remains to show that $G'(K) \ge 0$. Details of this argument are given in the appendix. ■

Theorem 3.2. *For aggregate loss A in the generalized collective risk model, the CV of the primary loss is no greater than the CV of the total loss; that is,*

$$CV_{A_p} \le CV_A.$$

Proof:

As in the proof of Theorem 3.1, we have

$$(CV_A)^2 = \left(\frac{\sigma_N}{\mu_N}\right)^2 + \frac{\mu_N E_\theta[\sigma^\theta_{X}] + E[N^2]Var_\theta[\mu^\theta_{X}]}{\mu_N^2\mu_{X}^2},$$

and

$$(CV_{A_p})^2 = \left(\frac{\sigma_N}{\mu_N}\right)^2 + \frac{\mu_N E_\theta[\sigma^\theta_{X_p}] + E[N^2]Var_\theta[\mu^\theta_{X_p}]}{\mu_N^2\mu_{X_p}^2}.$$

Therefore, it suffices to show that

$$\frac{\mu_N E_\theta[\sigma^\theta_{X_p}] + E[N^2]Var_\theta[\mu^\theta_{X_p}]}{\mu_{X_p}^2} \le \frac{\mu_N E_\theta[\sigma^\theta_{X}] + E[N^2]Var_\theta[\mu^\theta_{X}]}{\mu_{X}^2}.$$

To this end, we define a function

$$G(K) = \frac{\mu_N E_\theta[\sigma^\theta_{X_p}] + E[N^2]Var_\theta[\mu^\theta_{X_p}]}{\mu_{X_p}^2}.$$

In the appendix, we show that $G(K)$ is nondecreasing on R_+. Therefore, $G'(K) \ge 0$ on R_+, and the theorem is established. ■

In view of Theorem 3.1 and Theorem 3.2, we have the following conclusion.

Theorem 3.3. *For aggregate loss A in the generalized collective risk model, the CVs of the primary loss, the excess loss, and the total loss satisfy the following relationship:*

$$CV_{A_p} \le CV_A \le CV_{A_e}.$$

Theorem 3.3 generalizes Theorem 2.3. If the parameter space Θ is a singleton, that is, Θ contains only one element, then Theorem 3.1, Theorem 3.2, and Theorem 3.3 specialize to Theorem 2.1, Theorem 2.2, and Theorem 2.3, respectively. Therefore, the remarks at the end of Section 2 apply here too. This also means that we have given two different sets of proofs for Theorem 2.1, Theorem 2.2, and Theorem 2.3. However, for the collective risk model, the proofs in Section 2 are more straightforward than those in Section 3.

In addition, Theorem 3.3 has some potential applications in capital allocation. Suppose that an actuary needs to allocate capital to an employer's workers' compensation plan. The actuary decides to use the method of allocating capital by percentile layer, discussed extensively in Bodoff (2009) and Hong (2013),[4] but at the same time is concerned about the volatility of each layer of the future loss. Theorem 3.3

[4]To our knowledge, no consensus has been reached on the right approach to capital allocation. The literature documents many other methods; see Bauer and Zanjani (2013), Cummins (2000), D'Arcy (2011), and Venter (2004), and the references therein.

says that the layers of the excess loss are more volatile than the layers of the primary loss. Therefore, they might be more challenging to estimate accurately, and so the actuary might want to assign more capital to the layers of the excess loss.

Moreover, Theorem 3.3 might provide some new information for state regulators too. In view of Theorem 3.3, the excess loss is more volatile than the primary loss. Maybe more care needs to be exercised when regulators audit any calculation involving the excess loss.

4. Covariance and correlation coefficient between the primary loss and the excess loss

It is clear that the primary loss and the excess loss are related. Therefore, one would naturally look at the covariance and the correlation coefficient between them. The following theorem gives the covariance between the primary loss and the excess loss of a single claim.

Theorem 4.1. *Let X be a single claim, C be the state limit, and K be the split point. Then the primary loss and the excess loss are nonnegatively correlated, and the covariance, Cov(X_p, X_e), between them equals*

$$Cov(X_p, X_e) = \left(\int_0^K F(x)\,dx\right)\left(\int_K^C S(x)\,dx\right),$$

where F(x) and S(x) are the cumulative distribution function and survival function of X, respectively. Moreover, Cov(X_p,X_e) is bounded above by 1/4[$\Leftarrow_0^K F(x)dx$ + $\Leftarrow_K^C S(x)dx]^2$, *and the equality holds if and only if* $\Leftarrow_0^K F(x)dx = \Leftarrow_K^C S(x)dx$.

Proof:
Recall that

$$X_p = \begin{cases} X, & \text{if } X \le K; \\ K, & \text{if } X > K; \end{cases}$$

and that

$$X_e = \begin{cases} 0, & \text{if } X \le K; \\ X - K, & \text{if } K < X \le C; \\ C - K, & \text{if } X > C. \end{cases}$$

Therefore,

$$E[X_e] = \int_K^C (x-K)\,dF(x) + (C-K)S(C)$$
$$= \int_K^C S(x)\,dx;$$

$$E[X_p] = \int_0^K x\,dF(x) + KS(K) = \int_0^K S(x)\,dx;$$

and

$$E[X_e X_p] = \int_0^K 0 \cdot x\,dF(x) + \int_k^C (x-K)\cdot K\,dF(x)$$
$$+ \int_C^\infty (C-K)K\,dF(x)$$
$$= K\left[\int_K^C (x-K)\,dF(x) + (C-K)S(C)\right]$$
$$= KE[X_e].$$

It follows that

$$Cov(X_p, X_e) = E[X_p X_e] - E[X_p]E[X_e]$$
$$= KE[X_e] - \left(\int_0^K S(x)\,dx\right)E[X_e]$$
$$= \left(K - \int_0^K S(x)\,dx\right)E[X_e]$$
$$= \left(\int_0^K F(x)\,dx\right)\left(\int_K^C S(x)\,dx\right) \ge 0.$$

The second statement follows from the fact that for two nonnegative real numbers, a and b, $a + b \ge 2\sqrt{ab}$ and the equality holds if and only if $a = b$. ■

Theorem 4.1 says that X_p and X_e are positively correlated. That is, if the primary loss turns out to be larger than average, then it is likely that the excess loss will also be larger than average. This confirms the intuition that a larger-than-average X_e implies a larger-than-average X, which in turn implies a larger-than-average X_e. Moreover, Theorem 4.1 gives an explicit formula for $Cov(X_p, X_e)$, allowing actuaries to calculate the correlation coefficient between X_p and X_e as in Example 4.3. In other words, Theorem 4.1 makes it possible to measure the positive correlation between X_p and X_e on a numerical scale.

Since the primary loss and the excess loss clearly depend on K and C, it is interesting to see how K and C affect their CVs. A scrutiny of the proofs of Theorem 3.1 and Theorem 3.2 reveals that the following result holds.

Theorem 4.2. *For aggregate loss A in the generalized collective risk model, the CVs of the primary loss and the excess loss are both nondecreasing functions of K.*

In light of the previous discussion, it is natural to ask whether the covariance, $Cov(A_p, A_e)$, between the primary loss and the excess loss for aggregate loss A is also a monotonic function of K. The following example shows that the answer is negative.

Example 4.1. Suppose N is a degenerate random variable at the point 1; that is, we consider an individual claim X. We assume that X follows the exponential distribution with a hazard rate of $\lambda > 0$. Then the CDF, $F(x)$, and the survival function, $S(x)$, of X are given by

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0; \\ 0, & x \le 0; \end{cases}$$

and

$$S(x) = \begin{cases} e^{-\lambda x}, & x > 0; \\ 1, & x \le 0, \end{cases}$$

respectively. Assume $G(K) = Cov(A_p, A_e)$. Since N is degenerate at 1, $G(K) = Cov(X_p, X_e)$. It follows from Theorem 4.1 that

$$G(K) = \frac{1}{\lambda}\left(e^{-\lambda K} - e^{-\lambda C}\right)\left[K - \frac{1}{\lambda}\left(1 - e^{-\lambda K}\right)\right].$$

Table 4.1. Values of $G(K)$ for different values of K when $C = 10$

K	G(K)	K	G(K)	K	G(K)	K	G(K)
0.2	0.0153	1.2	0.1509	2.2	0.1452	3.2	0.0912
0.4	0.0471	1.4	0.1594	2.4	0.1352	3.4	0.0811
0.6	0.0871	1.6	0.1619	2.6	0.1243	3.6	0.0717
0.8	0.1120	1.8	0.1595	2.8	0.1131	3.8	0.0630
1.0	0.1353	2.0	0.1536	3.0	0.1020	4.0	0.0551

Take $\lambda = 1$. Then

$$G(K) = \left(e^{-K} - e^{-C}\right)\left(K - 1 + e^{-K}\right).$$

Table 4.1 gives some numerical values of $G(K)$ when $C = 10$, showing that $G(K)$ is not a monotonic function of K.

Remark. In particular, this example shows that the covariance, $Cov(X_{ip}, X_{je})$, between the primary loss and the excess loss of a single claim is not a monotonic function of K.

The fact that we used the dimensionless quantity *CV* in Sections 3 and 4 is important; otherwise, the conclusions in these sections may not hold. To illustrate this point, consider the following example.

Example 4.2. Suppose an individual claim, X, follows the exponential distribution with a hazard rate of $\lambda > 0$. Then $Var[X] = 1/\lambda^2$, and the following are true:

$$E[X_p] = \int_0^K x\lambda e^{-\lambda x}dx + KS(K) = \frac{1}{\lambda}\left(1 - e^{-\lambda K}\right),$$

$$E[X_p^2] = \int_0^K x^2\lambda e^{-\lambda x}dx + K^2S(K)$$
$$= \frac{2}{\lambda^2}\left(1 - e^{-\lambda K}\right) - \frac{2}{\lambda}Ke^{-\lambda K},$$

$$E[X_e] = \int_K^C (x - K)\lambda e^{-\lambda x}dx + (C - K)S(C)$$
$$= \frac{1}{\lambda}\left(e^{-\lambda K} - e^{-\lambda C}\right),$$

and

$$E\left[X_e^2\right]=\int_K^C (x-K)^2\,\lambda e^{-\lambda x}dx+(C-K)^2\,S(C)$$

$$=\frac{2}{\lambda}\left[\frac{1}{\lambda}\left(e^{-\lambda K}-e^{-\lambda C}\right)-(C-K)e^{-\lambda C}\right].$$

It follows that

$$Var\left[X_p\right]=\frac{1}{\lambda}\left[\frac{1}{\lambda}\left(1-e^{-2\lambda K}\right)-2Ke^{-\lambda K}\right];$$

$$Var\left[X_e\right]=\frac{1}{\lambda^2}\begin{bmatrix} e^{-\lambda K}\left(2-e^{-\lambda K}\right) \\ -e^{-\lambda C}\left(2+e^{-\lambda C}\right)+2e^{-\lambda(C+K)}\end{bmatrix} -\frac{2(C-K)}{\lambda}e^{-\lambda C}.$$

Therefore,

$$Var[X]-Var\left[X_p\right]=\frac{e^{-\lambda K}}{\lambda}\left(2K+\frac{1}{\lambda}e^{-\lambda K}\right)>0;$$

$$Var[X_e]-Var[X]$$

$$=\frac{1}{\lambda^2}\begin{bmatrix} e^{-\lambda K}\left(2-e^{-\lambda K}\right)-e^{-\lambda C}\left(2+e^{-\lambda C}\right) \\ +2e^{-\lambda(C+K)}-1\end{bmatrix} -\frac{2(C-K)}{\lambda}e^{-\lambda C}.$$

If $\lambda = 1$, $K = 1$, and $C = 4$, then

$$Var[X_e]-Var[X]=-0.533<0.$$

Indeed, if we look at the variance instead of the CV, then Theorem 4.1 implies that

$$Var[X]=Var\left[X_p+X_e\right]$$

$$=Var\left[X_p\right]+Var[X_e]+2Cov\left(X_p,X_e\right)$$

$$\geq Var\left[X_p\right]\text{ or }Var[X_e].$$

In view of the above discussion, we would like to see whether $Corr(A_p, A_e)$ is a monotonic function of K.

Table 4.2. Values of $G(K)$ for different values of K when $C = 10$

K	G(K)	K	G(K)	K	G(K)	K	G(K)
0.2	0.3335	1.2	0.4891	2.2	0.4495	3.2	0.3776
0.4	0.4160	1.4	0.4866	2.4	0.4364	3.4	0.3622
0.6	0.4565	1.6	0.4806	2.6	0.4225	3.6	0.3468
0.8	0.4772	1.8	0.4720	2.8	0.4079	3.8	0.3314
1.0	0.4868	2.0	0.4615	3.0	0.3929	4.0	0.3162

Since we have shown that CV_{A_e} and CV_{A_p} are both nondecreasing functions of K, one might expect that $Corr(A_p, A_e)$ is also a nondecreasing function of K. However, the next example shows that the answer is again negative.

Example 4.3. As in Example 4.1, we still consider the case in which N is degenerate at 1 and there is an individual loss, X, following the exponential distribution with a hazard rate of $\lambda = 1$. Recall that $G(K) = Corr(A_p, A_e) = Corr(X_p, X_e)$. It follows from Examples 4.1 and 4.2 that

$$G(K)=\frac{\left(e^{-K}-e^{-C}\right)\left(K-1+e^{-K}\right)}{\sqrt{\left(1-e^{-2K}-2Ke^{-K}\right)\begin{bmatrix} e^{-K}\left(2-e^{-K}\right) \\ -e^{-C}\left(2+2(C-K)+e^{-C}-2e^{-K}\right)\end{bmatrix}}}.$$

Table 4.2 gives some numerical values of $G(K)$ when $C = 10$; it shows that $G(K)$ is not a monotonic function of K.

Under the same assumption, we consider $Corr(A_p, A_e)$ as a function of the state limit, C; that is, $H(C) = Corr(A_p, A_e)$. Table 4.3 gives some numerical

Table 4.3. Values of $H(C)$ for different values of C when $K = 0.2$

C	H(C)	C	H(C)	C	H(C)	C	H(C)
0.5	0.6836	3.0	0.3768	5.5	0.3397	8.0	0.3342
1.0	0.5448	3.5	0.3632	6.0	0.3376	8.5	0.3339
1.5	0.4711	4.0	0.3538	6.5	0.3362	9.0	0.3337
2.0	0.4261	4.5	0.3473	7.0	0.3352	9.5	0.3336
2.5	0.3967	5.0	0.3428	7.5	0.3346	10.0	0.3335

values of $H(C)$ when $K = 0.2$; it suggests that $H(C)$ might be a nonincreasing function of C.

Since A_p is independent of C, intuitively one would expect $H(C)$ to be a monotonic function of C. The next theorem confirms this intuition.

Theorem 4.3. *For a single claim, X, the correlation coefficient, $Corr(X_p, X_e)$, between the primary loss and the excess loss is a nonincreasing function of C.*

Proof:

By Theorem 4.1, we have

$$Corr\left(X_p, X_e\right) = \frac{\left(\int_0^K F(x)dx\right)\left(\int_K^C S(x)dx\right)}{\sigma_{X_p}\sigma_{X_e}}.$$

Since the terms $\int_0^K F(x)dx$ and σ_{X_p} are both independent of C, and X_p and X_e are nonnegatively correlated, we obtain

$$H(C) = \frac{\left(\int_K^C S(x)dx\right)^2}{\sigma_{X_e}^2} = \frac{\mu_{X_e}^2}{\sigma_{X_e}^2},$$

showing that $H(C)$ is a nondecreasing function on R_+. We also have

$$\frac{\partial}{\partial C}\mu_{X_e}^2 = 2S(C)\mu_{X_e};$$

and

$$\frac{\partial}{\partial C}E\left[X_e^2\right] = 2(C-K)S(C).$$

It follows that

$$\frac{\partial}{\partial C}Var[X_e] = 2S(C)\left[(C-K)-\mu_{X_e}\right].$$

Therefore,

$$H'(C) = \frac{2S(C)\mu_{X_e}\sigma_{X_e}^2 - 2S(C)\left[(C-K)-\mu_{X_e}\right]\mu_{X_e}^2}{\sigma_{X_e}^4}$$

$$= \frac{2S(C)\mu_{X_e}\left[\sigma_{X_e}^2 - (C-K)\mu_{X_e} + \mu_{X_e}^2\right]}{\sigma_{X_e}^4}$$

$$= \frac{2S(C)\mu_{X_e}\left[E\left(X_e^2\right) - (C-K)\mu_{X_e}\right]}{\sigma_{X_e}^4}$$

$$= \frac{2S(C)\mu_{X_e}\left[\begin{array}{c}\int_K^C (x-K)^2 dF(x) \\ -(C-K)\int_K^C (x-K)dF(x)\end{array}\right]}{\sigma_{X_e}^4}$$

$$= \frac{2S(C)\mu_{X_e}\int_K^C (x-K)(x-C)dF(x)}{\sigma_{X_e}^4} \le 0. \quad \blacksquare$$

Theorem 4.3 says that the degree of correlation between X_p and X_e becomes smaller as C is getting larger. For a possible actuarial application, we consider NCCI's experience rating plan. Suppose the data for the excess loss is lost due to a recent cyber attack, but the data for the primary loss is still available. Now if the primary loss component is larger than average and the state limit is high, then actuaries know that there is a good chance the excess loss component will be larger than average. On the other hand, a quite low state limit gives more support to the possibility that the excess loss component will be larger than average.

As we pointed out earlier, two different claims, X_i and X_j, need not be independent in the generalized collective risk model; it is interesting to see whether X_{ip} and X_{je} are also nonnegatively correlated, that is, whether $Cov(X_{ip}, X_{je}) \ge 0$ holds. The next result shows that the answer is affirmative.

Theorem 4.4. *Let X_i and X_j be two distinct claims in the generalized collective risk model. Then the covariance between the primary loss of X_i and the excess loss of X_j is given by*

$$Cov\left(X_{ip}, X_{je}\right) = Cov_\theta\left[E\left(X_{ip}|\theta\right), E\left(X_{je}|\theta\right)\right]. \quad (4.1)$$

Moreover, $Cov(X_{ip}, X_{je}) \ge 0$; that is, X_{ip} and X_{je} are nonnegatively correlated.

Proof:

To establish Equation (4.1), we apply the conditional covariance formula to $Cov(X_{ip}, X_{je})$ and use the fact that the X_is are conditionally iid given θ to get

$$\begin{aligned} Cov(X_{ip}, X_{je}) &= E_\theta[Cov(X_{ip}, X_{je}|\theta)] \\ &\quad + Cov_\theta[E(X_{ip}|\theta), E(X_{je}|\theta)] \\ &= Cov_\theta[E(X_{ip}|\theta), E(X_{je}|\theta)]. \end{aligned}$$

To see the validity of the second statement, we notice that

$$\begin{aligned} &Cov_\theta[E(X_{ip}|\theta), E(X_{je}|\theta)] \\ &= E_\theta[E(X_{ip}|\theta)E(X_{je}|\theta)] - E[X_{ip}]E[X_{je}] \\ &= E_\theta[(E(X_{ip}|\theta) - E[X_{ip}])E(X_{je}|\theta)] \\ &\geq 0. \end{aligned}$$ ■

Theorem 4.4 also has some potential applications in actuarial science. Let us continue to consider the example discussed following Theorem 4.3—that is, the data for all past excess losses is lost due to a recent cyber attack, but the data for past primary losses is still available. Suppose the data shows that past primary losses were very large relative to their average. Then without obtaining any new data, an actuary may infer that the excess component of the next loss is likely to be large.

Example 4.1 shows that $Cov(A_p, A_e)$ is not a monotonic function of K. It is interesting to see from Theorem 4.1 that $Cov(X_p, X_e)$ is a nondecreasing function of C. Also note that $Cov(X_{ip}, X_{je})$ is not a monotonic function of K, as the next example shows.

Example 4.4. Suppose that, given the risk parameter θ, all claims $X_1, \ldots, X_n, \ldots$ are conditionally iid with a uniform distribution on $(0, \theta)$, and that θ follows a uniform distribution on $(1, a)$, where $0 \leq K \leq C \leq \theta$ and $a > 1$. Then we have

$$\begin{aligned} E[X_{ip}|\theta]) &= \int_0^K x\frac{1}{\theta}dx + K\left(\frac{\theta - K}{\theta}\right) \\ &= \frac{K(2\theta - K)}{2\theta}, \end{aligned}$$

and

$$\begin{aligned} E[X_{je}|\theta]) &= \int_K^C (x - K)\frac{1}{\theta}dx + (C - K)\left(\frac{\theta - C}{\theta}\right) \\ &= \frac{(C - K)(2\theta - K - C)}{2\theta}. \end{aligned}$$

Therefore,

$$\begin{aligned} E[X_{ip}] &= \int_\Theta \frac{K(2\theta - K)}{2\theta}dF(\theta) \\ &= \frac{K}{2(a-1)}\int_1^a \frac{(2\theta - K)}{\theta}d\theta \\ &= \frac{K}{2(a-1)}[2(a-1) - K\ln a] \\ &= K - \frac{K^2 \ln a}{2(a-1)}, \end{aligned}$$

$$\begin{aligned} E[X_{je}] &= \int_\Theta \frac{(C-K)(2\theta - K - C)}{2\theta}dF(\theta) \\ &= \frac{(C-K)}{2(a-1)}\int_1^a \frac{(2\theta - K - C)}{\theta}d\theta \\ &= \frac{(C-K)}{2(a-1)}[2(a-1) - (K+C)\ln a] \\ &= (C-K) - \frac{(C^2 - K^2)\ln a}{2(a-1)}, \end{aligned}$$

and

$$\begin{aligned} &E_\theta[E(X_{ip}|\theta)E(X_{je}|\theta)] \\ &= \int_\Theta \frac{K(C-K)(2\theta - K - C)(2\theta - K)}{4\theta^2}dF(\theta) \\ &= \frac{K(C-K)}{4(a-1)}\int_1^a \frac{(2\theta - K - C)(2\theta - K)}{\theta^2}d\theta \\ &= \frac{K(C-K)}{4(a-1)}\left[\begin{array}{l} 4(a-1) - (4K + 2C)\ln a \\ \quad + K(K+C)(1 - 1/a) \end{array}\right]. \end{aligned}$$

If $a = 5$, $C = 1$, and $G(K) = Cov(X_{ip}, X_{je})$, then, as Table 4.4 shows, $G(K)$ is not a monotonic function of K.

Table 4.4. Values of $G(K)$ for different values of K when $a = 5$ and $C = 1$

K	$G(K)$	K	$G(K)$	K	$G(K)$	K	$G(K)$
0.500	0.00179	0.625	0.00227	0.750	0.00234	0.875	0.00171
0.525	0.00190	0.650	0.00232	0.775	0.00229	0.900	0.00147
0.550	0.00201	0.675	0.00263	0.800	0.00219	0.925	0.00118
0.575	0.00211	0.700	0.00238	0.825	0.00207	0.950	0.00084
0.600	0.00219	0.725	0.00238	0.850	0.00191	0.975	0.00045

Table 4.5 gives some values of $Cov(X_{ip}, X_{je})$ for different values of C when $a = 5$ and $K = 0.25$.

Table 4.5 suggests that $H(C) = Cov(X_{ip}, X_{je})$ might be a nondecreasing function of C. The next result proves this fact.

Theorem 4.5. *Let X_i and X_j be two distinct claims in the generalized collective risk model. Then the covariance, $Cov(X_{ip}, X_{je})$, between the primary loss of X_i and the excess loss of X_j is a nonnegative, nondecreasing function of C.*

Proof:

If $H(C) = Cov(X_{ip}, X_{je})$, then in view of Theorem 4.1, we need only to show that $H(C)$ is nondecreasing. We have

$$H(C) = E_\theta\left[E\left(X_{ip}|\theta\right)E\left(X_{je}|\theta\right)\right] - E_\theta\left[E\left(X_{ip}|\theta\right)\right]E_\theta\left[E\left(X_{je}|\theta\right)\right]$$
$$= \int_\Theta E\left(X_{ip}|\theta\right)\left(\int_K^C S_{x|\theta}(x)\,dx\right)dF(\theta) - \left(\int_\Theta E\left(X_{ip}|\theta\right)dF(\theta)\right)\left(\int_\Theta\left(\int_K^C S_{x|\theta}(x)\,dx\right)dF(\theta)\right).$$

Since $E(X_{ip}|\theta)$ is independent of C,

$$H'(C) = \int_\Theta E\left(X_{ip}|\theta\right)S_{x|\theta}(C)\,dF(\theta) - \left(\int_\Theta E\left(X_{ip}|\theta\right)dF(\theta)\right)\left(\int_\Theta S_{x|\theta}(C)\,dF(\theta)\right)$$
$$= \int_\Theta E\left(X_{ip}|\theta\right)\left(S_{x|\theta}(C) - \int_\Theta S_{x|\theta}(C)\,dF(\theta)\right)dF(\theta) \geq 0.$$

Therefore, $H'(C)$ is nondecreasing function of C. ■

Examples 4.1 and 4.3 imply that neither $Cov(A_p, A_e)$ nor $Corr(A_p, A_e)$ is a monotonic function of K. However, the following theorem shows that $H(C) = Cov(A_p, A_e)$ is a nondecreasing function of C.

Theorem 4.6. *For aggregate loss A in the generalized collective risk model, the covariance, $Cov(A_p, A_e)$, between the primary loss and the excess loss is a nondecreasing function of C.*

Proof:

We have

$$Cov\left(A_p, A_e\right) = Cov\left(\sum_{i=1}^{N} X_{ip}, \sum_{j=1}^{N} X_{je}\right)$$

Table 4.5. Values of $H(C)$ for different values of C when $a = 5$ and $K = 0.25$

C	$H(C)$	C	$H(C)$	C	$H(C)$	C	$H(C)$
0.500	0.000112	0.625	0.000195	0.750	0.000298	0.875	0.000419
0.525	0.000127	0.650	0.000214	0.775	0.000320	0.900	0.000445
0.550	0.000143	0.675	0.000234	0.800	0.000344	0.925	0.000472
0.575	0.000160	0.700	0.000255	0.825	0.000368	0.950	0.000500
0.600	0.000177	0.725	0.000276	0.850	0.000393	0.975	0.000529

Table 4.6. Summary of the key results in Section 4

Function	Monotonicity	Reference
$G(K) = Cov(X_p, X_e)$	No	Example 4.1
$H(C) = Cov(X_p, X_e)$	Yes/nondecreasing	Theorem 4.1
$G(K) = Corr(X_p, X_e)$	No	Example 4.3
$H(C) = Corr(X_p, X_e)$	Yes/nonincreasing	Theorem 4.3
$G(K) = Cov(X_{ip}, X_{je})$	No	Example 4.4
$H(C) = Cov(X_{ip}, X_{je})$	Yes/nondecreasing	Theorem 4.5
$G(K) = Corr(X_{ip}, X_{je})$	Conjectured yes	Open problem
$H(C) = Corr(X_{ip}, X_{je})$	Conjectured yes	Open problem
$G(K) = Cov(A_p, A_e)$	No	Example 4.1
$H(C) = Cov(A_p, A_e)$	Yes/nondecreasing	Theorem 4.6
$G(K) = Corr(A_p, A_e)$	No	Example 4.3
$H(C) = Corr(A_p, A_e)$	Conjectured yes	Open problem

$$= \sum_{j=1}^{N}\sum_{i=1}^{N} Cov(X_{ip}, X_{je})$$

$$= \sum_{i=1}^{N} Cov(X_{ip}, X_{ie})$$

$$+ 2\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} Cov(X_{ip}, X_{je}).$$

Now the conclusion follows from Theorem 4.1 and Theorem 4.5. ■

Table 4.6 concludes this section by summarizing its key results.

5. Discussion

This paper has studied the properties of the primary loss and the excess loss in NCCI's split experience rating plan. It has shown that $CV_{A_p} \le CV_A \le CV_{A_e}$ in the collective risk model, generalizing a result obtained in the prior literature. Then it showed that the same conclusion holds in the generalized collective risk model. The last part of the paper investigated the covariance and correlation coefficient between the primary loss and the excess loss in the generalized collective risk model. The key results are summarized in Table 4.6.

In view of Theorem 4.5 and Theorem 4.6, we would like to know whether $Corr(X_{ip}, X_{je})$ is a nondecreasing function of both K and C. Our numerical analysis suggests that the answer might be affirmative. But a proof is elusive to us. Hence, we pose this open problem as a conjecture.

Conjecture 5.1. *Let X_i and X_j be two distinct claims in the generalized collective risk model. Then the correlation coefficient, $Corr(X_{ip}, X_{je})$, between the primary loss of X_i and the excess loss of X_j is a nonnegative, nondecreasing function of K. It is also a nonnegative, nondecreasing function of C.*

We also conjecture that $H(C) = Corr(A_p, A_e)$ is a nondecreasing function of C.

Conjecture 5.2. *For aggregate loss A in the generalized collective risk model, the correlation coefficient, $Corr(A_p, A_e)$, between the primary loss and the excess loss is a nondecreasing function of C.*

This paper has focused on developing some further results in the foundational theory of split credibility, though it has also discussed some potential applications of our results. Since we have limited knowledge, many potential applications may remain to be discovered. It is our hope that the results in this paper can stimulate some further work along this line.

Acknowledgments

Thanks are due to three anonymous reviewers for their thoughtful comments and suggestions, which led to several substantial improvements in this article. Thanks also go to Ryan Martin, who made the author realize that the usual setup of credibility theory entails an exchangeable model.

References

Bauer, D., and G. Zanjani, "Capital Allocation and Its Discontents," in G. Dionne, ed., *The Handbook of Insurance* (2nd ed.), pp. 863–880, Boston: Kluwer Academic Publishers, 2013.

Bodoff, N. M., "Capital Allocation by Percentile Layer," *Variance* 3:1, 2009, pp. 13–30.

Brockett, P. L., and J. R. Garven, "A Reexamination of the Relationship between Preferences and the Moment Orderings by Rational Risk-Averse Investors," *The Geneva Papers on Risk and Insurance Theory* 23, 1998, pp. 127–137.

Bühlmann, H., and A. Gisler, *A Course in Credibility Theory and Its Applications*, New York: Springer, 2005.

Cummins, J. D., "Allocation of Capital in the Insurance Industry," *Risk Management and Insurance Review* 3:1, 2000, pp. 7–27.

Cunningham, R., T. Herzog, and R. London, *Models for Quantifying Risk* (3rd ed.), Winsted, CT: Actex Publications, 2008

D'Arcy, S. P., "Capital Allocation in the Property-Liability Insurance Industry," *Variance* 5:2, 2011, pp. 141–157.

Gillam, W. R., "Workers Compensation Experience Rating: What Every Actuary Should Know," *Proceedings of the Casualty Actuarial Society* 79, 1992, pp. 215–239.

Heckman, P. E., and G. G. Meyers, "The Calculation of Aggregate Loss Distribution from Claim Severity and Claim Count Distributions," *Proceedings of the Casualty Actuarial Society* 70, 1983, pp. 22–61.

Hong, L., "Some Remarks on Capital Allocation by Percentile Layer," *European Actuarial Journal* 3, 2013, pp. 439–452.

Klugman, S. A., H. H. Panjer, and G. E. Willmot, *Loss Models: From Data to Decisions* (2nd ed.), Hoboken, NJ: Wiley, 2004.

Mahler, H. C., Discussion of "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *Proceedings of the Casualty Actuarial Society* 72, 1987, pp. 22–61.

Robbin, I., "The Theory of Split Credibility," *Variance* 7(1), 2013, pp. 29–60.

Schervish, M. J., *Theory of Statistics*, New York: Springer, 1995.

Venter, G. G., "Experience Rating—Equity and Predictive Accuracy," *NCCI Digest* 2(1), 1987, pp. 1–9.

Venter, G. G., "Capital Allocation Survey with Commentary," *North American Actuarial Journal* 8(2), 2004, pp. 96–107.

Appendix

From the proof of Theorem 3.1

We have

$$\frac{\partial}{\partial K}(\mu_{X_e}^2) = -2\mu_{X_e}S(K).$$

$$\begin{aligned}\frac{\partial}{\partial K}E_\theta\left[\sigma_{X_e}^\theta\right] &= E_\theta\left[\frac{\partial}{\partial K}\sigma_{X_e}^\theta\right]\\ &= E_\theta\left[\frac{\partial}{\partial K}\left(E\left[X_e^2|\theta\right]-\left(\mu_{X_e}^\theta\right)^2\right)\right]\\ &= E_\theta\left[-2\mu_{X_e}^\theta + 2\mu_{X_e}^\theta P(X>K|\theta)\right]\\ &= -2\mu_{X_e} + 2E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right],\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial K}Var_\theta\left[\mu_{X_e}^\theta\right] &= \frac{\partial}{\partial K}\left\{E_\theta\left[\left(\mu_{X_e}^\theta\right)^2\right]-\left(E_\theta\left[\mu_{X_e}^\theta\right]\right)^2\right\}\\ &= E_\theta\left[\frac{\partial}{\partial K}\left(\mu_{X_e}^\theta\right)^2\right]-\frac{\partial}{\partial K}\mu_{X_e}^2\\ &= E_\theta\left[-2\mu_{X_e}^\theta P(X>K|\theta)\right]+2\mu_{X_e}S(K)\\ &= 2\left\{\mu_{X_e}S(K)-E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]\right\}.\end{aligned} \tag{A.1}$$

The derivative $G'(K)$ equals

$$\frac{1}{\mu_{X_e}^4}\left\{\begin{bmatrix}-2\mu_N\mu_{X_e}+2\mu_N E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]\\ +E[N^2]\frac{\partial}{\partial K}Var_\theta\end{bmatrix}\mu_{X_e}^2 + 2\mu_{X_e}S(K)\begin{bmatrix}\mu_N E_\theta\left[\sigma_{X_e}^\theta\right]\\ +E[N^2]Var_\theta\left[\mu_{X_e}^\theta\right]\end{bmatrix}\right\}$$

$$=\frac{1}{\mu_{X_e}^3}\left\{\begin{bmatrix}-2\mu_N\mu_{X_e}+2\mu_N E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]\\ +E[N^2]\frac{\partial}{\partial K}Var_\theta\left[\mu_{X_e}^\theta\right]\end{bmatrix}\mu_{X_e} + 2S(K)\begin{pmatrix}\mu_N E_\theta\left[\sigma_{X_e}^\theta\right]\\ +E[N^2]Var_\theta\left[\mu_{X_e}^\theta\right]\end{pmatrix}\right\}.$$

To establish Equation (4), we will show that $G'(K)\geq 0$. Consider the following terms in the numerator of $G'(K)$:

$$\begin{aligned}&-2\mu_N\mu_{X_e}^2+2\mu_N S(K)E_\theta\left[\sigma_{X_e}^\theta\right]\\ &+2\mu_N\mu_{X_e}E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]\\ =&-2\mu_N\mu_{X_e}^2+2\mu_N S(K)E_\theta\left[E(X_e^2)-\left(\mu_{X_e}^\theta\right)^2|\theta\right]\\ &+2\mu_N\mu_{X_e}E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]\\ =&\,2\mu_N\left[S(K)E(X_e^2)-\mu_{X_e}^2\right]+2\mu_N\mu_{X_e}\\ &E_\theta\left[\mu_{X_e}^\theta P(X>K|\theta)\right]-2\mu_N S(K)E_\theta\left[\left(\mu_{X_e}^\theta\right)^2\right].\end{aligned}$$

The Cauchy-Schwarz inequality implies that $S(K)E[X_e^2]-\mu_{X_e}^2\geq 0$. Thus, we focus on the remaining two terms. Let us combine these two terms with the term $2S(K)E[N^2]Var_\theta[\mu^\theta_{X_e}]$ in the numerator of $G'(K)$. Since N takes nonnegative integer values, we have $N^2 \geq N$, which implies $E[N^2] \geq E[N]$. Therefore,

$$2\mu_N \mu_{X_e} E_\theta\left[\mu^\theta_{X_e} P(X > K|\theta)\right]$$
$$- 2\mu_N S(K) E_\theta\left[(\mu^\theta_{X_e})^2\right] + 2S(K)E[N^2]Var_\theta\left[\mu^\theta_{X_e}\right]$$
$$\geq 2\mu_N \left\{\begin{matrix} \mu_{X_e} E_\theta\left[\mu^\theta_{X_e} P(X > K|\theta)\right] - S(K)E_\theta\left[(\mu^\theta_{X_e})^2\right] \\ + S(K)Var_\theta\left[\mu^\theta_{X_e}\right] \end{matrix}\right\}.$$
$$= 2\mu_N \left[\begin{matrix} \mu_{X_e} E_\theta\left[\mu^\theta_{X_e} P(X > K|\theta)\right] - S(K)E_\theta\left[(\mu^\theta_{X_e})^2\right] \\ + S(K)E_\theta\left[(\mu^\theta_{X_e})^2\right] - S(K)\mu^2_{X_e} \end{matrix}\right]$$
$$= 2\mu_N \mu_{X_e}\left\{E_\theta\left[\mu^\theta_{X_e} P(X > K|\theta)\right] - S(K)\mu_{X_e}\right\}.$$

Up to this point, the only term in the numerator of $G'(K)$ that we have not used is $\mu_{X_e} E[N^2] \frac{\partial}{\partial K} Var_\theta\,[\mu^\theta_{X_e}]$, which is no less than $\mu_N\, \mu_{X_e} \frac{\partial}{\partial K} Var_\theta[\mu^\theta_{X_e}]$. In view of Equation (A.1), we see that the numerator of $G'(K) \geq 0$, implying that $G'(K) \geq 0$.

From the proof of Theorem 3.2

We have

$$\frac{\partial}{\partial K} E_\theta\left[\sigma^\theta_{X_p}\right] = E_\theta\left[\frac{\partial}{\partial K}\left(E(X_p^2|\theta) - (\mu^\theta_{X_p})^2\right)\right]$$
$$= E_\theta\left[2KP(X > K|\theta)\right]$$
$$- E_\theta\left[2\mu^\theta_{X_p} P(X > K|\theta)\right]$$
$$= 2KS(K) - 2E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right],$$

and

$$\frac{\partial}{\partial K} Var_\theta\left[\mu^\theta_{X_p}\right] = \frac{\partial}{\partial K} E_\theta\left[(\mu^\theta_{X_p})^2\right] - \frac{\partial}{\partial K}\left(E_\theta\left[\mu^\theta_{X_p}\right]\right)^2$$
$$= E_\theta\left[\frac{\partial}{\partial K}(\mu^\theta_{X_p})^2\right] - 2\mu_{X_p} S(K)$$
$$= 2E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] - 2\mu_{X_p} S(K).$$

Therefore, the derivative of $G(K)$ equals

$$\frac{1}{\mu^4_{X_p}}\left\{\begin{matrix} 2\mu^2_{X_p}\left[\begin{matrix} K\mu_N S(K) - \mu_N E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] \\ + E[N^2]E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] \\ - E[N^2]\mu_{X_p} S(K) \end{matrix}\right] \\ - 2\mu_{X_p} S(K)\left[\begin{matrix} \mu_N E_\theta\left[\sigma^\theta_{X_p}\right] \\ + E[N^2]Var_\theta\left[\mu^\theta_{X_p}\right] \end{matrix}\right] \end{matrix}\right\}$$

$$= \frac{2}{\mu^3_{X_p}}\left\{\begin{matrix} K\mu_N \mu_{X_p} S(K) - \mu_N \mu_{X_p} E_\theta \\ \left[\mu^\theta_{X_p} P(X > K|\theta)\right] \\ + \mu_{X_p} E[N^2]E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] \\ - E[N^2]\mu^2_{X_p} S(K) - \mu_N S(K)E_\theta\left[\sigma^\theta_{X_p}\right] \\ - S(K)E[N^2]Var_\theta\left[\mu^\theta_{X_p}\right] \end{matrix}\right\}.$$

Since $2/\mu_{X_p}{}^3 \geq 0$, it remains to show that the sum of

$$K\mu_N \mu_{X_p} S(K) - \mu_N \mu_{X_p} E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right]$$
$$+ \mu_{X_p} E[N^2]E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] - E[N^2]\mu^2_{X_p} S(K)$$
$$- \mu_N S(K)E_\theta\left[\sigma^\theta_{X_p}\right] - S(K)E[N^2]Var_\theta\left[\mu^\theta_{X_p}\right]$$

is nonnegative. The first, second, and the fifth terms sum to

$$\mu_N\left(K\mu_{X_p} S(K) - S(K)E\left[X_p^2\right]\right)$$
$$- \mu_N\left\{\mu_{X_p} E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] - S(K)E_\theta\left[(\mu^\theta_{X_p})^2\right]\right\}.$$

and the remaining three terms sum to

$$E[N^2]\mu_{X_p} E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] - E[N^2]\mu^2_{X_p} S(K)$$
$$- S(K)E[N^2]\left(E_\theta\left[(\mu^\theta_{X_p})^2\right] - \left(E_\theta\left[\mu^\theta_{X_p}\right]\right)^2\right)$$
$$= E[N^2]\left\{\begin{matrix} \mu_{X_p} E_\theta\left[\mu^\theta_{X_p} P(X > K|\theta)\right] \\ - S(K)E_\theta\left[(\mu^\theta_{X_p})^2\right] \end{matrix}\right\}.$$

As in the proof of Theorem 2.2, we know that $K\mu_{X_p} S(K) - S(K)E[X_p^2] \geq 0$. Hence, it remains to show that the sum of the other four terms is nonnegative. Indeed, these four terms add to

$$\left(E[N^2]-\mu_N\right)\begin{Bmatrix}\mu_{X_p}E_\theta\left[\mu_{X_p}^\theta P(X>K|\theta)\right]\\ -S(K)E_\theta\left[\left(\mu_{X_p}^\theta\right)^2\right]\end{Bmatrix}.$$

The same argument used in the proof of Theorem 3.1 shows that $E[N^2] - \mu_N \geq 0$. Also,

$$\begin{aligned}&\mu_{X_p}E_\theta\left[\mu_{X_p}^\theta P(X>K|\theta)\right]-S(K)E_\theta\left[\left(\mu_{X_p}^\theta\right)^2\right]\\ &=E_\theta\left[\mu_{X_p}^\theta\left[\mu_{X_p}P(X>K|\theta)-\mu_{X_p}^\theta S(K)\right]\right]\\ &\geq 0.\end{aligned}$$