On the properties of the primary loss and the excess loss in NCCI's experience rating plan

Liang Hong*

Abstract

The split credibility has been used in practice for several decades, though the foundational theory of split credibility was only investigated recently. This paper intends to study the properties of the primary loss and the excess loss in NCCI's split experience plan. We first re-visit the 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 where the frequency distribution is Poisson. We also show that the primary loss is less volatile than the total loss. Next, we show that the previously established ordering of the CV's 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.

*Liang Hong is a Fellow of the Society of Actuaries and an Associate Professor in the Department of Mathematics, Robert Morris University, 6001 University Boulevard, Moon Township, PA 15108, USA. Tel.: 412-397-4024. Email address: hong@rmu.edu.

1 Introduction

The American workers compensation system is one of the most successful social programs in the US. Many states in the US are currently using experience rating plan of National Council on Compensation Insurance (NCCI). In the current NCCI's experience rating plan, a cap, called the *state limit*, is first applied to each individual claim; next, a cutting 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; then two different credibilities are assigned to these two components. In the literature, such a practice is referred to as the *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 the 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 the CV of the excess loss in the collective risk model[1] when the frequency distribution is Poisson. We show that this conclusion holds for an arbitrary frequency distribution. In addition, we show that the CV of the primary loss is no greater than the CV of the total loss. This establishes the ordering of the CV's of the primary loss, the excess loss and the total loss.

Next, we point out that the usual credibility theory framework implies a model which is more general than 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 CV's 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; we establish several interesting results. For example, (i) neither the covariance nor the correlation coefficient between the

[1] To our knowledge, the terminology " collective risk model" has been defined differently in the prior literature. Our definition agrees with Definition 6.1 in Klugman et al (2004) and the definition in Chapter 13 of Cunningham et al (2008). The definition in Heckman and Meyers (1983) carries the same name, but it is slightly more general and is similar to the "generalized collective risk model" to be defined in Section 3.

primary loss and the excess loss of a claim is a monotone function of the split point; (ii) the covariance between the primary loss of a claim and the excess loss of a different claim is a non-decreasing function of the state limit; (iii) the covariance between the primary loss and the excess loss of a claim is a non-decreasing function of the state limit, but the correlation coefficient between the primary loss and the excess loss of a claim is a non-increasing function of the state limit; (iv) the primary loss and the excess loss of a single claim is non-negatively correlated; the primary loss of one claim and the excess loss of another is also non-negatively correlated. Table 6 summarizes the key results along this line.

We conclude our paper with two conjectures: (1) the correlation coefficient between the primary loss of one claim and the excess loss of another is a non-decreasing function of both the state limit and the split point; (2) the correlation coefficient between the primary loss and the excess loss of the aggregate loss is a non-decreasing function of the state limit.

The remainder of the paper is organized as follows. Section 2 establishes the fact the primary loss is less volatile than the total loss which 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.

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 coefficient of variation (CV) to measure the volatility of a random variable. Robbin (2013) noticed that an intuitive justification for the split credibility goes as follows. 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 Poisson, he proved that the CV of the total loss is no greater than the CV of the excess loss in the collective risk model. This result showed that the above-mentioned 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 too. 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 i-th claim ($i = 1, 2, ..., N$). We put $A = X_1 + X_2 + ... + X_N$, that is, A denotes the aggregate loss. We assume that all X_i's are independent and identically distributed (iid) with an absolutely continuous cumulative distribution function (CDF) $F(x)$ and a survival function $S(x)$. 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 $0 \leq K \leq C$. Then for the i-th claim, the primary loss and the excess loss are given by

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

and

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

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

$$A_p = X_{1p} + ... + X_{Np},$$

and

$$A_e = X_{1e} + ... + X_{Ne},$$

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. $\mu_X, \sigma_X, \mu_{X_p}, \sigma_{X_p}, \mu_{X_e}$ and σ_{X_e} should be interpreted similarly. For the aggregate loss A, the CV's 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 CV's 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 the 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,*

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

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

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

Proof: The square of the CV of the 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

$$\frac{\sigma_X^2}{\mu_X^2} \leq \frac{\sigma_{X_e}^2}{\mu_{X_e}^2}. \tag{1}$$

To this end, we write Equation (1) as

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

which is equivalent to

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

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

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

and

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

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

$$
\mu_{X_e}^2 \leq S(K)E[X_e^2]. \tag{3}
$$

To establish Equation (3), we write μ_X^2 as

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

where

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

and apply the Cauchy-Schwarz inequality. □

Theorem 2.2 *For the 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,*

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

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

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

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

$$\frac{E[X_p^2]}{\mu_{X_p}^2} \le \frac{E[X^2]}{\mu_X^2}.$$

To this end, we put

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

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 only need to show that $G(K)$ is non-decreasing on R_+. We have

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

and

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

It follows that

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

where the last inequality follows from the fact that

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

This completes the proof. □

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

Theorem 2.3 *For the aggregate loss A in the collective risk model, the CV's 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, the CV's 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 CV's 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's experience rating plan splits each the aggregate loss into two components with one being less volatile and the other being more volatile. Therefore, the above-mentioned justification is incorrect regardless of the distribution of the claim frequency. For the correct justification, see Robbin (2013).

In the collective risk model, all claim severity random variables are assumed to be independent identically distribution (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 the case where the claim severity random variables are conditionally iid and see whether the key results in this section will still hold. 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 the credibility theory, each of the observed claims $X_1, ..., X_n, ...$ 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 $\theta \in \Theta$, $X_1, ..., X_n, ...$ are assumed to be iid. Though $X_1, ..., X_n, ...$ 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

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

in such a model; see, for example, Bühlmann and Gisler (2005)[2]. Therefore, we consider the following model: claim severity random variables $X_1, ..., X_n, ...$ are conditional iid given the risk parameter θ and the frequency random variable N are independent of all the X_i's. To distinguish this model from the collective risk model in Section 2, we call this model the *generalized collective risk model*[3]. Then it is natural to ask whether the conclusion of Theorem 2.3 still holds in the generalized collective risk model. The rest of the section is devoted to showing that the answer is affirmative. To lighten the notation,

[2] Technically, $X_1, ..., X_n, ...$ are said to be *exchangeable*. That is, for every finite collection $X_{i_1}, ..., X_{i_n}$ every permutation of $(X_{i_1}, ..., X_{i_n})$ has the same joint distribution. This follows from the DeFinetti Theorem; see, for example, Theorem 1.49 of Schervish (1995). In words, the basic framework of credibility theory entails an exchangeable model.

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

we will use μ_X^θ and σ_X^θ to denote $E[X \mid \theta]$ and $Var(X \mid \theta)$, respectively. $\mu_{X_e}^\theta, \sigma_{X_e}^\theta, \mu_{X_p}^\theta, \sigma_{X_p}^\theta$ are to be interpreted in a similar manner.

Theorem 3.1 *For the 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*

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

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

$$\mu_A = E_N[E[A \mid N]] = E[NE[X]] = E[N]E[X] = \mu_N \mu_X.$$

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

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

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

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

It follows from the previous two displays that

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

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

$$\begin{aligned} \mu_{A_e} &= \mu_N \mu_{X_e}; \\ \sigma_{A_e}^2 &= \sigma_N^2 \mu_{X_e}^2 + \mu_N E_\theta[\sigma_{X_e}^\theta] + E[N^2] Var_\theta[\mu_{X_e}^\theta]; \\ (CV_{A_e})^2 &= \left(\frac{\sigma_N}{\mu_N}\right)^2 + \frac{\mu_N E_\theta[\sigma_{X_e}^\theta] + E[N^2] Var_\theta[\mu_{X_e}^\theta]}{\mu_N^2 \mu_{X_e}^2}. \end{aligned}$$

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

$$\frac{\mu_N E_\theta[\sigma_X^\theta] + E[N^2]Var_\theta[\mu_X^\theta]}{\mu_X^2} \leq \frac{\mu_N E_\theta[\sigma_{X_e}^\theta] + E[N^2]Var_\theta[\mu_{X_e}^\theta]}{\mu_{X_e}^2}. \quad (4)$$

To this end, we put

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

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

Theorem 3.2 *For the 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,*

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

Proof: As in the proof of Theorem 3.1, we have

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

and

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

Therefore, it suffices to show that

$$\frac{\mu_N E_\theta[\sigma_{X_p}^\theta] + E[N^2]Var_\theta[\mu_{X_p}^\theta]}{\mu_{X_p}^2} \leq \frac{\mu_N E_\theta[\sigma_X^\theta] + E[N^2]Var_\theta[\mu_X^\theta]}{\mu_X^2}.$$

To this end, we define a function

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

In the Appendix, we show that $G(K)$ is non-decreasing on R_+. Therefore, $G'(K) \geq 0$ 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 the aggregate loss A in the generalized collective risk model, the CV's of the primary loss, the excess loss and the total loss satisfy the following relationship:*

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

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, those remarks at the end of Section 2 apply here too. This also means that we have given two different 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.

In addition, Theorem 3.3 have 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 percentile layer discussed extensively in Bodoff (2009) and Hong (2013)[4]. At the same time, he or she is concerned about the volatility of each layer of the future loss. Then Theorem 3.3 says that the layers of the excess loss are more volatility than the layers of the primary loss. Therefore, they might be more challenging to estimate accurately. In view of this, 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 the 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 involved 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 correlation coefficient between them. The following theorem gives the covariance between the primary loss and the excess loss of a single claim.

[4]To our knowledge, no consensus has been reached on the right approach to it 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.

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 non-negatively 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 $1/4[\int_0^K F(x)dx + \int_K^C S(x)dx]^2$ and the equality holds if and only if $\int_0^K F(x)dx = \int_K^C S(x)dx$.

Proof: Recall that

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

and

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

Therefore,

$$\begin{aligned} E[X_e] &= \int_K^C (x-K)dF(x) + (C-K)S(C) = \int_K^C S(x)dx; \\ E[X_p] &= \int_0^K xdF(x) + KS(K) = \int_0^K S(x)dx; \\ E[X_eX_p] &= \int_0^K 0 \cdot xdF(x) + \int_K^C (x-K)\cdot KdF(x) + \int_C^\infty (C-K)KdF(x) \\ &= K\left[\int_K^C (x-K)dF(x) + (C-K)S(C)\right] = KE[X_e]. \end{aligned}$$

It follows that

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

The second statement follows from the fact that for two non-negative real number a and b we have $a + b \geq 2\sqrt{ab}$, where the equality holds if and only if $a = b$. $\square$

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 that a larger-than-average X_e. Moreover, Theorem 4.1 gives an explicit formula for $Cov(X_p, X_e)$. This allows actuaries to calculate the correlation coefficient between X_p and X_e as in Example 4.3. In 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 the K and C, it is interesting to see how K and C affect the CV's of them. A scrutiny of the proofs of Theorem 3.1 and Theorem 3.2 reveals that the following result holds.

Theorem 4.2 *For the aggregate loss A in the generalized collective risk model, the CV's of the primary loss and the excess loss are both non-decreasing functions of K.*

Therefore, it is natural to ask whether the covariance $Cov(A_p, A_e)$ between the primary loss and the excess loss for the aggregate loss A is also a monotone function of K. The next example shows that the answer is negative.

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

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

and

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

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

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

Take $\lambda = 1$. Then

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

Table 1 gives some numerical values of $G(K)$ when $C = 10$; it shows that $G(K)$ is not a monotone function of K.

Table 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

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 monotone 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 $\lambda > 0$. Then $Var[X] = 1/\lambda^2$. Also, we have

$$\begin{aligned}
E[X_p] &= \int_0^K x\lambda e^{-\lambda x}dx + KS(K) = \frac{1}{\lambda}(1 - e^{-\lambda K}), \\
E[X_p^2] &= \int_0^K x^2\lambda e^{-\lambda x}dx + K^2 S(K) = \frac{2}{\lambda^2}(1 - e^{-\lambda K}) - \frac{2}{\lambda}Ke^{-\lambda K}, \\
E[X_e] &= \int_K^C (x-K)\lambda e^{-\lambda x}dx + (C-K)S(C) = \frac{1}{\lambda}\left(e^{-\lambda K} - e^{-\lambda C}\right), \\
E[X_e^2] &= \int_K^C (x-K)^2\lambda e^{-\lambda x}dx + (C-K)^2 S(C) \\
&= \frac{2}{\lambda}\left[\frac{1}{\lambda}\left(e^{-\lambda K} - e^{-\lambda C}\right) - (C-K)e^{-\lambda C}\right].
\end{aligned}$$

It follows that

$$\begin{aligned}
Var[X_p] &= \frac{1}{\lambda}\left[\frac{1}{\lambda}(1 - e^{-2\lambda K}) - 2Ke^{-\lambda K}\right]; \\
Var[X_e] &= \frac{1}{\lambda^2}\left[e^{-\lambda K}(2 - e^{-\lambda K}) - e^{-\lambda C}(2 + e^{-\lambda C}) + 2e^{-\lambda(C+K)}\right] \\
&\quad -\frac{2(C-K)}{\lambda}e^{-\lambda C}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
Var[X] - Var[X_p] &= \frac{e^{-\lambda K}}{\lambda}\left(2K + \frac{1}{\lambda}e^{-\lambda K}\right) > 0; \\
Var[X_e] - Var[X] &= \frac{1}{\lambda^2}\left[e^{-\lambda K}(2 - e^{-\lambda K}) - e^{-\lambda C}(2 + e^{-\lambda C}) + 2e^{-\lambda(C+K)} - 1\right] \\
&\quad -\frac{2(C-K)}{\lambda}e^{-\lambda C}.
\end{aligned}$$

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

$$\begin{aligned}
Var[X] &= Var[X_p + X_e] \\
&= Var[X_p] + Var[X_e] + 2Cov(X_p, X_e) \\
&\geq Var[X_p] \text{ or } Var[X_e].
\end{aligned}$$

In view of the above discussion, we would like to see whether $Corr(A_p, A_e)$ is a monotone function of K. Since we have showed that CV_{A_e} and CV_{A_p} are both non-decreasing functions of K, one might expect that $Corr(A_p, A_e)$ is also a non-decreasing 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 where N is degenerate at 1 and an individual loss X following the exponential distribution with a hazard rate $\lambda = 1$. Put $G(K) = Corr(A_p, A_e) = Corr(X_p, X_e)$. It follows from Examples 4.1 and 4.2 that

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

Table 2 gives some numerical values of $G(K)$ when $C = 10$; it shows that $G(K)$ is not a monotone function of K.

Table 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

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 3 gives some numerical values of $H(C)$ when $K = 0.2$; it suggests that $H(C)$ might be a non-increasing function of C. Since A_p is independent of C, intuitively one would expect $H(C)$ to a monotone 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 non-increasing function of C.*

Table 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

Proof: By Theorem 4.1, we have

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

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 non-negatively correlated, we put

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

and show that $H(C)$ is a non-decreasing function on R_+. We have

$$\begin{aligned} \frac{\partial}{\partial C}\mu_{X_e}^2 &= 2S(C)\mu_{X_e}; \\ \frac{\partial}{\partial C}E[X_e^2] &= 2(C-K)S(C). \end{aligned}$$

It follows that

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

Therefore,

$$
\begin{aligned}
H'(C) &= \frac{2S(C)\mu_{X_e}\sigma^2_{X_e} - 2S(C)[(C-K)-\mu_{X_e}]\mu^2_{X_e}}{\sigma^4_{X_e}} \\
&= \frac{2S(C)\mu_{X_e}\left[\sigma^2_{X_e} - (C-K)\mu_{X_e} + \mu^2_{X_e}\right]}{\sigma^4_{X_e}} \\
&= \frac{2S(C)\mu_{X_e}\left[E(X_e^2) - (C-K)\mu_{X_e}\right]}{\sigma^4_{X_e}} \\
&= \frac{2S(C)\mu_{X_e}\left[\int_K^C (x-K)^2 dF(x) - (C-K)\int_K^C (x-K)dF(x)\right]}{\sigma^4_{X_e}} \\
&= \frac{2S(C)\mu_{X_e}\int_K^C (x-K)(x-C)dF(x)}{\sigma^4_{X_e}} \le 0.
\end{aligned}
$$

□

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 the average and the state limit is high, then actuaries know that there is a good chance that the excess loss component is less likely to be larger than the average. On the other hand, a quite low state limit gives more support to the possibility that the excess loss component is larger than the 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 non-negatively 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*

$$Cov(X_{ip}, X_{je}) = Cov_\theta\left[E(X_{ip} \mid \theta), E(X_{je} \mid \theta)\right]. \tag{5}$$

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

Proof: To establish Equation (5), we apply the conditional covariance formula to $Cov(X_{ip}, X_{je})$ and use the fact that X_i's are conditionally iid given θ to get

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

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

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

□

Theorem 4.4 also has some potential applications in actuarial science. Still consider the example 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 the past primary losses is still available. Suppose the data shows that the past primary losses were very large relative to their average. Then without obtaining any new data, an actuary may infer that the excess component for the next loss is likely to be large.

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

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

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

and

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

Therefore,

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

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

and

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

Table 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

Take $a = 5$ and $C = 1$ and put $G(K) = Cov(X_{ip}, X_{je})$. Then Table 4 shows that $G(K)$ is not a monotone function of K.

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

Table 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 5 suggests that $H(C) = Cov(X_{ip}, X_{je})$ might be a non-decreasing 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 non-negative non-decreasing function of C.*

Proof: Put $H(C) = Cov(X_{ip}, X_{je})$. In view of Theorem 4.1, we only need to show that $H(C)$ is non-decreasing. We have

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

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

$$
\begin{aligned}
H'(C) &= \int_\Theta E(X_{ip} \mid \theta)\ S_{x|\theta}(C)dF(\theta) \\
&\quad -\left(\int_\Theta E(X_{ip} \mid \theta)dF(\theta)\right)\left(\int_\Theta S_{x|\theta}(C)dF(\theta)\right) \\
&= \int_\Theta E(X_{ip} \mid \theta)\left(\ S_{x|\theta}(C) - \int_\Theta S_{x|\theta}(C)dF(\theta)\right) dF(\theta) \\
&\geq 0.
\end{aligned}
$$

Therefore, $H'(C)$ is non-decreasing 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 monotone function of K. However, the following theorem shows that $H(C) = Cov(A_p, A_e)$ is a non-decreasing function of C.

Theorem 4.6 *For the 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 non-decreasing function of C.*

Proof: We have

$$
\begin{aligned}
Cov(A_p, A_e) &= Cov\left(\sum_{i=1}^N X_{ip}, \sum_{j=1}^N X_{je}\right) \\
&= \sum_{j=1}^N \sum_{i=1}^N Cov(X_{ip}, X_{je}) \\
&= \sum_{i=}^N Cov(X_{ip}, X_{ie}) + 2\sum_{i=1}^{N-1} \sum_{j=i+1}^N Cov(X_{ip}, X_{je}).
\end{aligned}
$$

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

We conclude this section by presenting the following Table 6 that summarizes the key results obtained in this section.

Table 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/non-decreasing	Theorem 4.1
$G(K) = Corr(X_p, X_e)$	No	Example 4.3
$H(C) = Corr(X_p, X_e)$	Yes/non-increasing	Theorem 4.3
$G(K) = Cov(X_{ip}, X_{je})$	No	Example 4.4
$H(C) = Cov(X_{ip}, X_{je})$	Yes/non-decreasing	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/non-decreasing	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

5 Discussion

In this paper, we studied the properties of the primary loss and the excess loss in NCCI's split experience rating plan. We showed that $CV_{A_p} \leq CV_A \leq CV_{A_e}$ in the collective risk model, generalizing a result in the prior literature. Then we showed that the same conclusion holds in the generalized collective risk model. In the last part of the paper, we investigate 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 6.

In view of Theorem 4.5 and Theorem 4.6, we would like to know whether $Corr(X_{ip}, X_{je})$ is a non-decreasing 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 non-negative non-decreasing function of K. It is also a non-negative non-decreasing function C.*

We also conjecture that $H(C) = Corr(A_p, A_e)$ is a non-decreasing functions of C.

Conjecture 5.2 *For the 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 non-decreasing function of C.*

We focused on developing some further results in the foundational theory of the split credibility, though we also discussed some potential applications of our results. Since we have limited knowledge, there can be many potential applications 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 the three anonymous reviewers for their thoughtful comments and suggestions which led to several substantial improvements in this article. The author also thanks Ryan Martin who made the author realize that the usual setup of credibility theory entails an exchangeable model.

References

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

Bodoff, N.M. (2009). Capital allocation by percentile layer, *Variance*, 3(1), pp. 13–30.

Brockett, P.L. and Garven, J.R. (1998). A Reexaminiation of the relstionship between preferences and the moment orderings by rational risk-averse investors, *The Geneva Papers on Risk and Insurance Theory*, 23, pp. 127–137.

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

Cummins, J.D. (2001). Allocation of capital in the insurance industry, *Risk Management and Insurance Review*, 3(1), pp. 7–27.

Cunningham, R., Herzog, T. and London, R. (2008). *Models for Quantifying Risk*, Third Edition, Winsted: Actex Publication.

D'Arcy, S. P. (2011). Capital allocation in the property-liability insurance industry, *Variance*, 5(2), pp. 141–157.

Gillam, W.R. (1992). Workers compensation experience rating: what every actuary should know, *Proceedings of the Casualty Actuarial Society*, 79, pp. 215-239.

Heckman, P.E. and Meyers, G.G. (1983). The calculation of aggregate loss distribution from claim severity and claim count distributions, *Proceedings of the Casualty Actuarial Society*, pp. 22-61.

Hong, L. (2013). Some remarks on capital allocation by percentile layer, *European Actuarial Journal*, 3, pp. 439–452.

Klugman, S. A., Panjer, H. H. and Willmot, G. E. (2004). *Loss Models: from Data to Decisions*, Second Edition, Hoboken: Wiley.

Mahler, H.C. (1987). Discussion of " The calculation of aggregate loss distributions from claim severity and claim count distributions, " *Proceedings of the Casualty Actuarial Society*, Vol. LXXII, pp. 22-61.

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

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

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

Venter, G.G. (2004). Capital allocation survey with commentary, *North American Actuarial Journal*, 8(2), pp. 96–107.

Appendix

From the Proof of Theorem 3.1

We have

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

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

and

$$\begin{aligned}
\frac{\partial}{\partial K} Var_\theta[\mu_{X_e}^\theta] &= \frac{\partial}{\partial K}\left\{E_\theta\left[(\mu_{X_e}^\theta)^2\right] - \left(E_\theta[\mu_{X_e}^\theta]\right)^2\right\} \\
&= E_\theta\left[\frac{\partial}{\partial K}\left(\mu_{X_e}^\theta\right)^2\right] - \frac{\partial}{\partial K}\mu_{X_e}^2 \\
&= E_\theta\left[-2\mu_{X_e}^\theta P(X > K \mid \theta)\right] + 2\mu_{X_e} S(K) \\
&= 2\left\{\mu_{X_e} S(K) - E_\theta\left[\mu_{X_e}^\theta P(X > K \mid \theta)\right]\right\}. \qquad (6)
\end{aligned}$$

The derivative $G'(K)$ equals

$$
\begin{aligned}
& \frac{1}{\mu_{X_e}^4}\Bigg\{\Bigg[-2\mu_N\mu_{X_e}+2\mu_N E_\theta[\mu_{X_e}^\theta P(X>K\mid\theta)] \\
& +E[N^2]\frac{\partial}{\partial K}Var_\theta[\mu_{X_e}^\theta]\Bigg]\mu_{X_e}^2 \\
& +2\mu_{X_e}S(K)\Bigg[\mu_N E_\theta[\sigma_{X_e}^\theta]+E[N^2]Var_\theta[\mu_{X_e}^\theta]\Bigg]\Bigg\} \\
= \; & \frac{1}{\mu_{X_e}^3}\Bigg\{\Bigg[-2\mu_N\mu_{X_e}+2\mu_N E_\theta[\mu_{X_e}^\theta P(X>K\mid\theta)] \\
& +E[N^2]\frac{\partial}{\partial K}Var_\theta[\mu_{X_e}^\theta]\Bigg]\mu_{X_e} \\
& +2S(K)\left(\mu_N E_\theta[\sigma_{X_e}^\theta]+E[N^2]Var_\theta[\mu_{X_e}^\theta]\right)\Bigg\}.
\end{aligned}
$$

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

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

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

Therefore,

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

Up to this point, the only term in the numerator of $G'(K)$ 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 (6), we see that the numerator of $G'(K) \geq 0$, implying $G'(K) \geq 0$.

From the Proof of Theorem 3.2

We have

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

and

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

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

$$\begin{aligned}
& \frac{1}{\mu_{X_p}^4}\Bigg\{2\mu_{X_p}^2\Big[K\mu_N S(K) - \mu_N E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] \\
& +E[N^2]E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] - E[N^2]\mu_{X_p} S(K)\Big] \\
& -2\mu_{X_p} S(K)\Big[\mu_N E_\theta[\sigma_{X_p}^\theta] + E[N^2]Var_\theta[\mu_{X_p}^\theta]\Big]\Bigg\} \\
= \; & \frac{2}{\mu_{X_p}^3}\Bigg\{K\mu_N\mu_{X_p} S(K) - \mu_N\mu_{X_p} E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] \\
& +\mu_{X_p} E[N^2]E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] - E[N^2]\mu_{X_p}^2 S(K) \\
& -\mu_N S(K) E_\theta[\sigma_{X_p}^\theta] - S(K)E[N^2]Var_\theta[\mu_{X_p}^\theta]\Bigg\}.
\end{aligned}$$

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

$$\begin{aligned}
& K\mu_N\mu_{X_p} S(K) - \mu_N\mu_{X_p} E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] \\
& +\mu_{X_p} E[N^2]E_\theta[\mu_{X_p}^\theta P(X > K \mid \theta)] - E[N^2]\mu_{X_p}^2 S(K) \\
& -\mu_N S(K) E_\theta[\sigma_{X_p}^\theta] - S(K)E[N^2]Var_\theta[\mu_{X_p}^\theta]
\end{aligned}$$

is non-negative. The first, second and the fifth terms sum to

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

and the remaining three terms sum to

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

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

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

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

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