

# Advancements in Common Shock Modeling

## ABSTRACT

In this paper we rigorously investigate The Common Shock Model for correlating insurance losses. In addition, we develop additional theory which describes how The Common Shock Model, or Contagion model, can be incorporated within a larger set of distributions. We also address the issue of calibrating contagion models to empirical data. To this end, we propose several procedures for calibrating Contagion models using real-world Industry data. Finally, we demonstrate the efficacy, and efficiency, of these calibration procedures by calibrating aggregate loss models, which incorporate Contagion. Further, these case-studies illustrate the power of Contagion modeling by demonstrating how the introduction of Contagion can correct for the short-comings of traditional Collective Risk Modeling.

## Acknowledgements:

The authors wish to recognize the contribution to the following parties:

Benedict M Escoto PhD, FCAS, *Aon Benfield* Chicago, Illinois, for being willing to share his knowledge on contagion modeling, and supporting calibration.

Steven B White, FCAS & Chief Actuary, *Guy Carpenter*, LLC. for sharing his *contagious* enthusiasm for Contagion modeling with us.

Stu Sadwin, FCAS MAAA SVP & Chief Corporate Actuary, TMNA Services, LLC, for his invaluable support on the model calibration.

## 1 Introduction

The Common Shock Model, which was introduced by *Glenn Meyers* (Meyers, 1997), and is also called *Contagion Modeling*, *Steve White* (White, 2012), is a recent approach to modeling the systematic effect of changing insurance industry and general economic environments on losses. These effects include factors specific to the insurance industry, such as the so-called *underwriting cycle*, as well as cycles in the overall business environment, such as economic recessions and legal changes. Often the effect of such an environmental change is to cause losses to exhibit similar behavior both within, as well as across, companies or lines of business. This behavior often takes the form of a correlation of losses within, or across, companies or lines of businesses. The Common Shock Model, or Contagion model method attempts to capture, and account for, the effects of such an influential environment, which is referred to as a *contagious environment*. These models accomplish this by determining common factors which can be incorporated, separately, into both the frequency and severity distributions of effected losses. For the sake of simplicity, we use the single name Contagion model to refer to The Common Shock Model.

This paper is structured as follows; Section 2 briefly describes the traditional collective risk modeling paradigm, which is ubiquitous within general insurance modeling. Next, Sections 3, and 4, will explain how contagion modeling is built upon the collective risk modeling framework. Specifically, it will be shown how both the frequency and severity components of the collective risk model can be broadened so that the effects of contagious environments can be modeled. Further, Section 3 describes how frequency contagion can be incorporated within the common families of frequency distributions, and Section 4 describes how the effect of a contagious environment can be modeled within the severity component of the collective risk model. Next, Section 5 describes the aggregate loss distribution under contagion modeling. Section 6 describes the original work on the calibration of contagion models using empirical data, as completed by the research team at Tokio Marine Technologies. Lastly, Section 7 provides several case-studies on the application of contagion modeling to general insurance losses. In particular, Section 7 illustrates the application of the most direct calibration scheme, from Section 6, to real-world insurance data, and makes clear the potential deficiency of traditional collective risk modeling as applied to aggregate annual layered losses.

## 2 Collective Risk Modeling

Collective Risk Modeling refers to the bifurcation of an aggregate loss random variable (RV) into two RV's, one modeling the Frequency, or number of claims, and one modeling the Severity, or size of claims. Let  $N$  denote the number of claims arising from policies, in a given time period. Let  $X_1$  denote the amount of the first claim, and  $X_2$  the amount of the second claim, and so on. In the collective risk model, the random sum;

$$S = X_1 + X_2 + \cdots + X_N \quad (2.1)$$

represents the aggregate claims generated by the portfolio over the period under study. The number of claims  $N$  is a random variable, and is associated with the frequency of claims. The individual claims  $X_1, X_2, \dots$  are also random variables, and are said to measure the severity of claims. There are two fundamental assumptions in the collective risk model. The first is that  $X_1, X_2, \dots$  are independent and identically distributed (i.i.d.) random variables, and the second is that the random variables  $N, X_1, X_2, \dots$  are mutually independent.

For the collective risk model, the expected value of aggregate claims is the product of the expected individual claim amount and the expected number of claims;  $E(S) = E(N)E(X)$ , while the variance of aggregate claims is the sum of two components where the first is attributed to the variability of individual claim amounts, and the other to the variability of the number of claims:

$$Var(S) = E(N) \cdot Var(X) + [E(x)^2] \cdot Var(N) \quad (2.2)$$

The formulas for the mean and variance of aggregate losses under the collective risk model are fundamentally dependent on the assumption that  $X_1, X_2, \dots$  are *i.i.d.*, and that they are independent of the frequency RV,  $N$ . In Section 3, and Section 4, it will be explained how contagion modeling is built upon the collective risk modeling framework. Specifically, it will be shown how both the frequency and severity components of the collective risk model can be broadened so that the effects of contagious environments can be modeled.

### 3 Frequency contagion

Let  $N_i$  represent the random variable for the number of claims in the  $i$ -th line of business. Each  $N_i$  may be modeled using a common frequency distributions, such as the Poisson, Negative Binomial, or Binomial distribution. If each of the  $i$  lines of business falls within the same contagious environment, a common contagion RV can be incorporated within each of the  $i$  frequency distributions. In this section, it is explained how this can be accomplished for the three most common frequency distributions in the general insurance industry; the Poisson, Negative Binomial, and Binomial distributions. The incorporation of contagion within the Poisson and Negative Binomial distributions were originally put forth by Meyers, however the incorporation of contagion within the Binomial distribution, in the current work, is original.

#### 3.1 Poisson contagion model

The Poisson distribution is the most popular distribution used for modeling count processes in the insurance industry. This is not only because the claim count generation process can often be closely approximated with a Poisson distribution, but also because of the great flexibility of the Poisson distribution. The Poisson distribution is a discrete, single-parameter distribution, of the following form:

$$Pr(N = n) = \frac{\lambda^n e^{-\lambda}}{n!} \quad \text{for } n = 1, 2, \dots \quad \text{and } \lambda > 0 \quad (3.1.1)$$

To denote that a RV  $N$  follows a poisson distribution with parameter  $\lambda$  the following abbreviation will be used in the following;  $N \sim \text{Poisson}(\lambda)$ . Since Poisson frequency contagion has been documented in the literature, the basics of frequency contagion using the Poisson distribution is included in Appendix 9.1.

Let  $N_i$ , for  $i = 1, 2, \dots, K$  be  $K$  claim count RV's. The RV's can represent  $K$  lines of business, or claim counts from  $K$  different companies, etc. For the remainder of this paper we will refer to such a group of RV's as coming from  $K$  lines of business, with the understanding that the group of RV's could represent other situations, as well, and all results will still apply. Further, assume that each of the  $K$  lines of business is subject to, or influenced by, the same contagious environment. Also, assume that it has been determined that a Poisson distribution is suitable for modeling the counts from each of the  $K$  lines of business. It is important to note that the parameter of the Poisson distribution for the  $i^{th}$  line of business can be determined by fitting the model to only the data for the  $i^{th}$  line of business, and hence is not influenced by the value of the parameters for the other lines. We can describe this:

$$N_i \sim \text{Poisson}(\lambda_i) \quad \text{for } i = 1, 2, \dots, K \quad (3.1.2)$$

Now, to incorporate the effects of the contagious environment on each of the  $K$  lines of business, an additional RV, denoted by  $C$ , is utilized. Specifically,  $C$  is incorporated within each of the  $K$  Poisson RV's through multiplication with the Poisson parameter  $\lambda_i$ . Explicitly, the RV,  $C$ , is incorporated within each Poisson distribution, as follows:

$$Pr(N_i = n | C) = \frac{(C \cdot \lambda_i)^n e^{-C \cdot \lambda_i}}{n!} \quad \text{for } n = 1, 2, \dots \quad \text{and } \lambda_i > 0, \quad C \geq 0 \quad (3.1.3)$$

which can be denoted, more succinctly, by:  $N_i | C \sim \text{Poisson}(C \cdot \lambda_i)$ . It is important to reinforce that  $C$  is, itself, a RV, and as the notation implies, each  $N_i$  depends, or is conditional upon,  $C$ . The corresponding marginal distribution, or predictive distribution, of  $N$  is:

$$Pr(N_i = n) = \int_0^\infty pr(N_i = n | C) g(C) dC \quad (3.1.4)$$

Hence, to induce frequency contagion within a set of Poisson RV's, the mean of each of Poisson,  $\lambda_i$ , is multiplied by the *same* RV,  $C$ . We call  $C$  the *contagion random variable*. More will be said, shortly, about the application of contagion across several claim-count distributions. First, the impact of incorporating contagion *within* a single claim count distribution is investigated more closely.

One of the goals of contagion modeling is to construct parsimonious models that capture the effects which are induced by a contagious environment, while at the same time, producing

models that are faithful to the empirical data. For this reason, it is desired that the constructed models match the data, as much as possible. At minimum, it is desirable for the mean of a contagion model to match the mean produced by the corresponding traditional frequency model. However, since the Poisson contagion model is formed by multiplying the mean of the Poisson distribution by  $C$ , without restriction on the distribution of  $C$  we would have that:

$$E(N_i) = E_C[E_{N_i}(N_i | C)] = E_C[C \cdot \lambda_i] = E_C[C] \cdot \lambda_i \quad (3.1.5)$$

Hence, in order to ensure that  $E(N_i) = \lambda_i$ , it is necessary to place some restrictions on the distribution of  $C$ . For this this purpose, requiring that  $E(C) = 1$  will ensure that:

$$E(N_i) = \lambda_i \quad (3.1.6)$$

So far, this is the only restriction on the distribution of  $C$ . However, the motivation for including  $C$ , in the first place, is to facilitate the modeling of effects produced by the presence of a contagious environment. As a result, additional characteristics of the distribution of  $C$  must be considered. First, note that if  $C$  equals one, with probability 1, then the distribution of  $N_i$  is unaffected by the inclusion of  $C$ . This is equivalent to viewing  $C$  as having a distribution with mean one, and variance 0. Conversely,  $C$  can be imagined as having a distribution with mean one, but infinite variance. This can be viewed as representing an infinitely chaotic environment. In this case, one can imagine that the number of claims, from one period to another, will vary wildly. In this case, the variance of the claim counts will be extremely large.

Now, consider the additional restriction that the variance of  $C$  is equal to some constant value, which is denoted by  $c$ , i.e.  $Var[C] = c$ . With this additional restriction, the variance of the Poisson claim count process,  $N_i$ , under contagion, can be calculated as:

$$Var(N_i) = Var_C[E[N_i|C]] + E_C[Var[N_i|C]] = Var_C[C\lambda_i] + E_C[C\lambda_i] = \lambda_i(1 + c \cdot \lambda_i) \quad (3.1.7)$$

Hence, it can be seen that when the variance of  $C$  equals 0, namely  $c = 0$ , the variance of  $N_i$  will equal that of the original Poisson distribution. Alternatively, as the variance of  $C$  grows, the variance of  $N_i$  will also grow. In practice, it is the value of the contagion parameter  $c$ , rather than the variance of  $N_i$ , which the practitioner will adjust, or deduce from the data. From Equation (3.1.7), it can be seen that the variance of  $N_i$  increases linearly as a function of  $c$ , and quadratically as a function of  $\lambda_i$ . It is important to reiterate that the above results hold *regardless of the distribution of  $C$* . It is only required that:

- 1) the distribution of  $C$  has positive support (the domain is a subset of the positive reals).
- 2) the expected value of  $C$  is 1, and,
- 3) the variance of  $C$  is  $c$ .

To denote that  $C$  can take on any distribution, which satisfies these requirements, we use the following notation:

$$C \sim \text{Randist}(E[C] = 1, Var[C] = c) \quad (3.1.8)$$

However, the Poisson frequency contagion model conforms to the common conjugate prior relationships, from elementary Bayesian analysis. In particular, if  $C$  follows a gamma distribution, then  $N_i$  will have a negative binomial distribution with parameters  $r = 1/c$  and  $p = 1/(1 + c\lambda_i)$  such that  $E(N_i) = \lambda_i$  and  $Var(N_i) = \lambda_i(1 + c\lambda)$ . By the same token, if  $C$  follows an arbitrary distribution, it is likely that the exact distributional form of  $N_i$  will be unknown. However, the mean and variance will have the above form, and the contagion parameter will effect the model in same way.

So far, only frequency contagion models, when claim counts are modeled with a Poisson distribution, have been described. Moreover, only a single claim count distribution has been considered. In practice, what is often of primary interest when considering a contagious environment is the *shared impact*, or *correlation*, among effected claim count distributions. The Poisson contagion model, however, is amenable to this type of analysis. By simply utilizing the same contagion RV,  $C$ , across several claim counts distributions, the shared impact, or correlation, between the count distributions can be measured. This *across* lines-of-business, or *across* company, frequency contagion can be depicted as:

$$\begin{cases} N_1|C \sim \text{Poisson}(C \cdot \lambda_1) \\ N_2|C \sim \text{Poisson}(C \cdot \lambda_2) \\ \vdots \\ N_K|C \sim \text{Poisson}(C \cdot \lambda_K) \end{cases} \quad \text{where } C \sim \text{Randist}(E[C] = 1, Var[C] = c) \quad (3.1.9)$$

Again, it is important to note that each  $N_i$  can have a different mean,  $\lambda_i$ , but as long as the same contagion RV,  $C$ , is used for each  $N_i$ , it can be seen that:

$$E[N_i N_j] = E_C [E[N_i N_j | C]] = \lambda_i \lambda_j E_C [C^2] \quad \text{for } 1 \leq i, j \leq K \text{ and } i \neq j \quad (3.1.10)$$

Thus, the covariance between the claim counts from two lines of business, where  $1 \leq i, j \leq K$ , and  $i \neq j$ , is:

$$Cov(N_i, N_j) = E[N_i N_j] - E[N_i]E[N_j] = \lambda_i \lambda_j (1 + c) - \lambda_i \lambda_j = c \lambda_i \lambda_j \quad (3.1.11)$$

Hence, for  $1 \leq i, j \leq K$ , and  $i \neq j$  the correlation between  $N_i, N_j$  is:

$$\rho_{N_i, N_j} = \frac{Cov(N_i, N_j)}{\sigma_{N_i} \sigma_{N_j}} = \frac{c \lambda_i \lambda_j}{\sqrt{\lambda_i(1 + c\lambda_i)} \sqrt{\lambda_j(1 + c\lambda_j)}} = \sqrt{\frac{c\lambda_i}{1 + c\lambda_i}} \sqrt{\frac{c\lambda_j}{1 + c\lambda_j}} \quad (3.1.12)$$

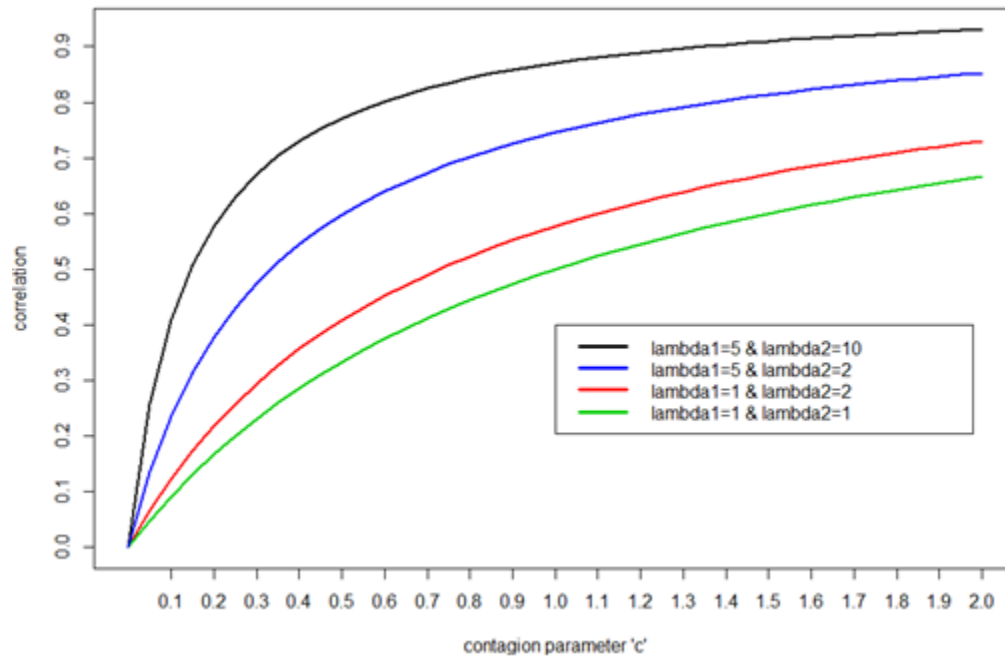
It is worth noting that, even though the same contagion RV,  $C$ , is used for each  $N_i$ , the induced correlation between any two  $N_i$  &  $N_j$  may be different from the correlation between any other

pair. As can be seen from Equation (3.1.12), this is because the correlation between  $N_i$  &  $N_j$  depends on the value of the means;  $\lambda_i$  and  $\lambda_j$ . Equation (3.1.12) also shows that the level of the contagion parameter,  $c$ , impacts the correlation between  $N_i$  &  $N_j$  in a way that makes intuitive sense. Namely:

- 1) As  $c \rightarrow 0$ , which represents a *weak*, or *absent*, contagious environment,  $\rho_{N_i, N_j} \rightarrow 0$ .
- 2) As  $c \rightarrow \infty$ , which represents a *strong* contagious environment,  $\rho_{N_i, N_j} \rightarrow 1$ .

Figure 1 depicts the pair-wise correlation between  $N_i$  &  $N_j$  as a function of  $c$ , and also for different combination of means,  $\lambda_i$  and  $\lambda_j$ .

**Figure 1: Correlation (across-lines) vs.  $c$ , and per level of Poisson frequency means.**



Note, from Figure 1 that larger values of  $\lambda$  result in higher correlation. This also makes intuitive sense, as there is less likely to be a strong relationship between the few claims produced each year by two frequency distributions with small means. Alternatively, when each frequency distribution predicts a large number of claims each year, there is more opportunity for such a relationship to exist.



### 3.2 Negative binomial contagion

The concept of frequency contagion was likely conceived with the Poisson distribution in mind. The fact that the Poisson is a single-parameter distribution, with parameter equaling the mean, as well as the well-known conjugate relationship with Gamma distribution, make the extension from traditional frequency modeling to frequency *contagion* modeling seem almost obvious. However, the application of frequency contagion modeling is not limited to the Poisson distribution. If the practical application of frequency contagion were limited to only one distribution, then the argument for contagion modeling would be less persuasive. In (Meyers, 1997) contagion was incorporated within frequency RV's that followed a Negative Binomial distribution, but only indirectly. This was accomplished by starting with the Poisson distribution and then exploiting the relationship between Negative Binomial distribution and the Poisson distribution under contagion. In this section, we explicitly illustrate how to incorporate a contagion RV within a Negative Binomial distribution, and hence make it clear how to start with a Negative Binomial distribution and subsequently introduce contagion. In Appendix 9.2 we present the basics of the Negative Binomial distribution, including some of its various parameterizations.

To illustrate the incorporation of contagion within a Negative Binomial distribution, we first begin with the  $r, \beta$  parameterization of the Negative Binomial distribution:

$$\Pr(N = n) = \binom{r + n - 1}{n} \left( \frac{1}{1 + \beta} \right)^r \left( \frac{\beta}{1 + \beta} \right)^n \quad \text{for } n \geq 0 \text{ and } \beta > 0, r \in \mathbb{Z}^+ \quad (3.2.1)$$

Then the contagion RV,  $C$ , can simply be multiplied by the  $\beta$  parameter, analogous to the way the contagion RV is incorporated within the Poisson distribution by multiplying  $\lambda$  by  $C$ . Hence, we can denote the Negative Binomial distribution, under contagion, by:

$$N|C \sim \text{NegBin}(r, C\beta) \quad \text{with } C \sim \text{Randist}(E[C] = 1, \text{Var}[C] = c) \quad \text{where } C \in \mathbb{Z}^+ \quad (3.2.2)$$

This,  $r$  and  $\beta$ , parameterization of the Negative Binomial clearly illustrates the incorporation of the contagion RV, however, the derivations of many results become more clear when an alternative parameterization of the Negative Binomial is used. The Negative binomial distribution can also be parameterized using the mean of  $\lambda$ , and a dispersion parameter,  $\gamma$ , which is related to variance-to-mean ratio:  $1 + \gamma\lambda$ , i.e.

$$N \sim \text{NegBin}(\lambda, \gamma) \quad (3.2.3)$$

This parameterization is related to the  $r$  and  $\beta$ , parameterization as follows:

$$N \sim \text{NegBin}(r = 1/\gamma, \beta = \gamma\lambda) \quad (3.2.4)$$

And under this  $\lambda, \gamma$  parameterization the *mean*,  $\lambda$ , is multiplied by the contagion RV, in order to induce contagion:

$$N|C \sim \text{NegBin}(C\lambda, \gamma) \quad (3.2.5)$$

Notice that if  $C$  takes the constant value of 1, with probability 1, then  $N$  will be an ordinary Negative Binomial distribution with mean  $\lambda$  and variance  $\lambda(1 + \gamma\lambda)$ , or  $r\beta$  and  $r\beta(1 + \beta)$ . As was done for the Poisson, we now investigate what restrictions need to be placed on the distribution of the RV,  $C$ . First, note that since  $C$  is multiplied by  $\lambda$ , we again require that  $C$  take on only positive values. Next, using the same *tower-property* of conditional expected values, we see that:

$$E(N) = E_C[E_N(N | C)] = E_C[C \cdot \lambda] = E_C[C] \cdot \lambda \quad (3.2.6)$$

Hence, by requiring that  $C$  has expected value of 1, or  $E[C] = 1$ , we have:

$$E(N) = \lambda \quad (3.2.7)$$

Further, if we assume that  $\text{Var}[C] = c$ , as shown in the Appendix 9.3, then:

$$\text{Var}(N) = \lambda(1 + \lambda(c + c\gamma + \gamma)) \quad (3.2.8)$$

Equation (3.2.8) shows that if the contagion parameter is 0, then the distribution of  $N$  reverts to the original Negative Binomial, whereas as  $c$  grows, the variance increases. Hence, as was the case for the Poisson distribution, the only necessary restrictions on the distribution of  $C$  are that  $C$  takes on only positive values, and  $E[C] = 1$ , and  $\text{Var}[C] = c$ , where  $c$  is the contagion parameter. Again, we denote this by:

$$C \sim \text{RanDist}(E[C] = 1, \text{Var}[C] = c) \quad (3.2.9)$$

Now, as before, assume that there are  $K$  lines of business, each of which is subject to, or influenced by, the same contagious environment. Also, assume that parameters of the best-fitting Negative Binomial have been fit to the data for each line of business. Then, as before, we simply use the same contagion RV,  $C$ , within each of the  $K$  Negative Binomial distributions, in order to model this contagious environment:

$$\begin{cases} N_1 | C \sim \text{NegBin}(C\lambda_1, \gamma_1) \\ N_2 | C \sim \text{NegBin}(C\lambda_2, \gamma_2) \\ \vdots \\ N_K | C \sim \text{NegBin}(C\lambda_K, \gamma_K) \end{cases} \quad \text{where } C \sim \text{RanDist}(E[C] = 1, \text{Var}[C] = c) \quad (3.2.10)$$

Again, it is important to note that each  $N_i$  can have different parameters,  $\lambda_i$  and  $\gamma_i$ , or  $r_i$  and  $\beta_i$ , after fitting to the respective data. However, as long as the same contagion RV,  $C$ , is used for each  $N_i$ , each  $N_i$  will be related. Further, Appendix 9.4 shows that the covariance between the number of claims from any two lines of business, subject to the same contagious environment, can be calculated as:

$$\text{Cov}(N_i, N_j) = c\lambda_i\lambda_j, \text{ for } 1 \leq i, j \leq K \text{ and } i \neq j. \quad (3.2.11)$$

And Appendix 9.5 shows that the resulting correlation between  $N_i$  and  $N_j$ , for  $1 \leq i, j \leq K$  and  $i \neq j$ , is:

$$\rho_{N_i, N_j} = \sqrt{\frac{c\lambda_i}{1 + \lambda_i(c + c\gamma_i + \gamma_i)}} \sqrt{\frac{c\lambda_j}{1 + \lambda_j(c + c\gamma_j + \gamma_j)}} \quad (3.2.12)$$

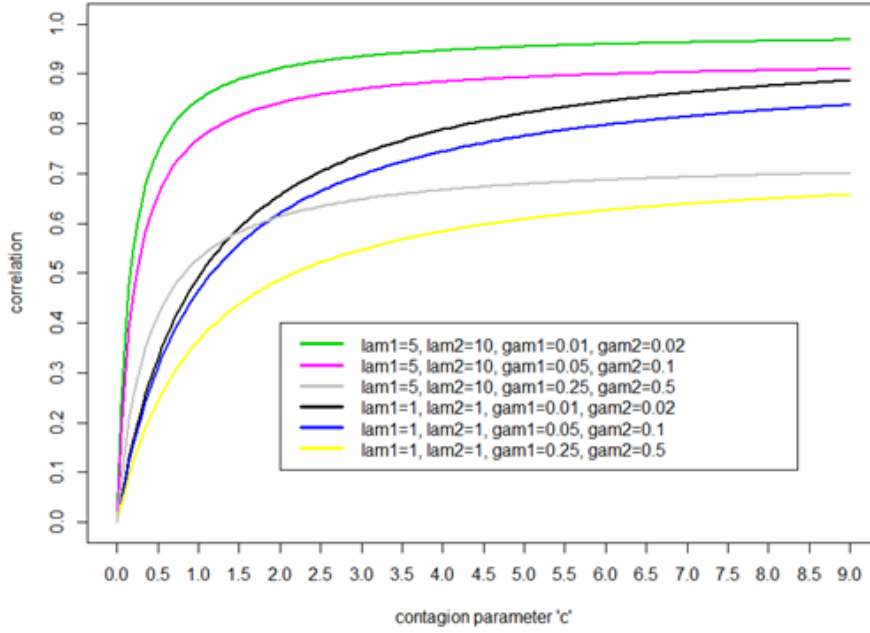
As can be seen from Equation (3.2.12), the correlation between  $N_i$  &  $N_j$  depends on the parameter values of each distribution;  $\gamma_i$  and  $\gamma_j$ , which again, may cause unique correlations to be induced for each  $N_i$  &  $N_j$ , even though the same contagion parameter,  $c$ , is used. Equation (3.2.12) also shows that the level of the contagion parameter,  $c$ , impacts the correlation between  $N_i$  &  $N_j$  in a way that makes intuitive sense. Namely:

- 1) As  $c \rightarrow 0$ , which represents a *weak*, or *absent*, contagious environment,  $\rho_{N_i, N_j} \rightarrow 0$ .
- 2) As  $c \rightarrow \infty$ , which represents a *strong* contagious environment,  $\rho_{N_i, N_j} \rightarrow \sqrt{\frac{1}{1+\gamma_1}} \sqrt{\frac{1}{1+\gamma_2}}$ .

Note that, in general, the dispersion parameters,  $\gamma_i$ , will tend to be small for insurance data. So even though the correlation may not equal 1 as  $c \rightarrow \infty$ , it should attain a value reasonably close to 1.

Figure 2, below, depicts the pair-wise correlation between  $N_i$  &  $N_j$  as a function of  $c$ , and also for different combination of the parameters,  $\lambda_i, \gamma_i$  and  $\lambda_j, \gamma_j$ . In Figure 2 it can be seen that, for large  $c$ , the correlation between  $N_i$  and  $N_j$  can be made close to 1, in all cases except when the dispersion parameters,  $\gamma$ , takes on the largest values in this example; 0.5 and 0.25. As the name implies, the dispersion parameter represent the amount of additional variance over and above the mean of a distribution. Since variance is a measure of uncertainty, when there is a large amount of uncertainty in two RV, there is correspondingly less certainty in the relationship between the two variables. Hence, Equation (3.2.12) for the correlation between two Negative Binomial distributions, under contagion, makes intuitive sense. Namely, when the dispersion parameters are high, there is less correlation. There is also a trade-off between the dispersion parameters and the means,  $\lambda$ , in the correlation formula. Even when the dispersion parameters are relatively large, sufficiently large values of the means,  $\lambda_i$ , will offset the deleterious effect on the correlation caused by large dispersion parameters.

Figure 2: Correlation (across-lines) vs.  $c$ , and per level of Negative Binomial parameters.



### 3.3 Binomial-Beta contagion

The three most common frequency distributions used within actuarial science are the Poisson, Negative Binomial, and the Binomial. For this reason, we also present a method for incorporating contagion within a Binomial distribution, as well. To the authors' knowledge, the literature on contagion modeling does not include a viable method for incorporating contagion within the Binomial distribution, at least in such a way that positive correlations can be achieved.

If the empirical variance of claim count data is significantly less than the corresponding mean, a Binomial distribution is often fit to the data. The probability density for a Binomial distribution with parameters  $n_i$  and  $p_i$  can be written as:

$$N \sim \text{Binomial}(n, p) = \binom{n}{N} p^N (1-p)^{n-N} \quad \text{where } 0 \leq N \leq n \text{ and } n \in \mathbb{Z}^+, 0 < p \leq 1 \quad (3.3.1)$$

When only dealing with a single Binomial distribution, contagion can be incorporated within the Binomial distribution by considering the  $p$  parameter to be a RV, and then using a suitable distribution for  $p$ . Assume, first, that the best-fitting Binomial distribution has been determined, and results in the parameter estimates,  $\hat{p}$  and  $\hat{n}$ . Next, we assume that  $N \sim \text{Binomial}(\hat{n}, p)$ , where only  $p$  is considered to be a RV, and the parameter  $n$  is evaluated at  $\hat{n}$ . Next, consider a prior distribution for  $p$  of the following form, with contagion parameter  $c$ :

$$p \sim \text{Beta} \left[ \alpha = \frac{1}{c}, \beta = \left( \frac{1}{c} \right) \frac{1 - \hat{p}}{\hat{p}} \right] \quad (3.3.2)$$

Note that both  $\hat{p}$  and  $c$  are constants, in this Beta distribution. Using this distribution for  $p$  we can check that the resulting posterior predictive distribution for  $N$  has the desired properties. In particular, that the mean is unaffected, while at the same time, larger values of  $c$  result in a larger variance. In Appendix 9.6 it is shown that:

$$E(N) = E_p[E_N(N | p)] = \hat{n} \cdot \hat{p} \quad (3.3.3)$$

Thus the mean of the Beta-Binomial mixture equals the mean of the original, best-fitting, Binomial. Next, in Appendix 9.7 it is shown that the variance of the Beta-Binomial mixture is:

$$Var(N) = \hat{n}\hat{p}(1 - \hat{p}) \cdot \frac{1 + \hat{n}\hat{p}c}{1 + \hat{p}c} \quad (3.3.4)$$

Notice that when  $c \rightarrow 0$ ,  $Var(N) \rightarrow \hat{n}\hat{p}(1 - \hat{p})$ , and as  $c \rightarrow \infty \Rightarrow Var(N) \rightarrow \hat{n}[\hat{n}\hat{p}(1 - \hat{p})] > \hat{n}\hat{p}(1 - \hat{p})$ .

Now, let  $N_i$  be the number of claims for the  $i^{th}$  line of business, out of  $K$  lines of business. If we believe that the claim counts for each of the  $K$  lines of business are subject to the same contagious environment, we can apply contagion across the  $K$  Binomial frequency distributions. First, for each of the  $K$  lines of business, assume that the best-fitting Binomial distribution has been determined. Denote the parameter estimates which specify these best-fitting Binomial distributions by;  $\{\hat{n}_i, \hat{p}_i \mid 1 \leq i \leq K\}$ . Next, in order to induce correlation across the  $K$  lines of business, the same prior must be used for each of the  $K$  Binomial distributions. However, due to the form of the prior distribution on  $p$ , namely Equation (3.3.2), without additional adjustment, the mean of  $N_i$  would equal  $\hat{n}_i\hat{p}$ , rather than  $\hat{n}_i, \hat{p}_i$ . This is not ideal, as the means of the Binomials, under contagion, will not all equal the means of the best-fitting Binomials, unless  $\hat{p} = \hat{p}_1 = \dots = \hat{p}_K$ . One approach to correct for this is to define:

$$\hat{p}^* = \max(\hat{p}_1, \hat{p}_2, \dots, \hat{p}_n) \quad (3.3.5)$$

Then, adjust each of the  $K$  Binomial frequency distributions by multiplying the parameter,  $p$ , of each distribution by the constant ratio  $\hat{p}_i/\hat{p}^*$ , as such:

$$N_i | p \sim Bin\left(\hat{n}_i, \left(\frac{\hat{p}_i}{\hat{p}^*}\right)p\right) \quad (3.3.6)$$

Here, the motivation for defining  $\hat{p}^*$  as the max of all the  $\hat{p}_i$  is clear, as this results in having  $0 \leq (\hat{p}_i/\hat{p}^*)p \leq 1$  over the support of the prior distribution for  $p$ . Lastly, the full, cross-line, Binomial contagion model can be formed by using the, above, adjusted Binomial distributions for each line, in conjunction with the same Beta distribution prior, with parameters;  $\alpha = \frac{1}{c}$  and  $\beta = \frac{1}{c}\left(\frac{1-\hat{p}}{\hat{p}}\right)$ :

$$\begin{cases} N_1|p \sim \text{Bin}\left(\hat{n}_1, \left(\frac{\hat{p}_1}{\hat{p}^*}\right)p\right) \\ N_2|p \sim \text{Bin}\left(\hat{n}_2, \left(\frac{\hat{p}_2}{\hat{p}^*}\right)p\right) \\ \vdots \\ N_K|p \sim \text{Bin}\left(\hat{n}_K, \left(\frac{\hat{p}_K}{\hat{p}^*}\right)p\right) \end{cases} \quad \text{where} \quad p \sim \text{Beta}\left(\alpha = \frac{1}{c}, \beta = \frac{1}{c}\left(\frac{1-\hat{p}^*}{\hat{p}^*}\right)\right) \quad (3.3.7)$$

We assert that this framework results in posterior predictive distributions, for each  $N_i$ , which have the desired properties, while at the same time can capture a reasonable range of dependencies between each of  $N_i$ . First, we verify the assertions regarding the marginal distribution of each  $N_i$ , and then the assertions regarding the joint distribution of the  $N_i$ .

Analogous to the computation for a single Binomial RV, in Appendix 9.8 it is shown that, for  $1 \leq i \leq K$ :

$$E(N_i) = E_p[E_{N_i}(N_i | p)] = \hat{n}_i \cdot \hat{p}_i \quad (3.3.8)$$

Hence, each marginal distribution under the cross-line Binomial contagion model has a mean equal to the, possibly unique, mean of the best-fitting Binomial distribution to the data from that line.

Next we verify the characteristics of the variance. In Appendix 9.9 it is shown that, for  $1 \leq i \leq K$ , the variance for a single Binomial RV is:

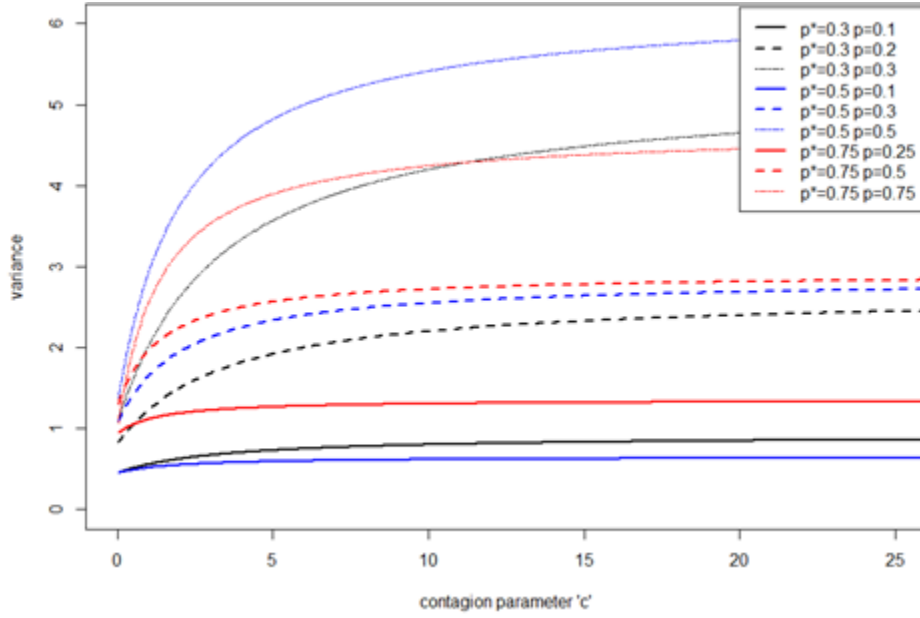
$$\text{Var}(N) = \frac{\hat{n}_i \hat{p}_i (1 - \hat{p}_i) + c \hat{p}^* \left[ \hat{n}_i \hat{p}_i \left(1 - \frac{\hat{p}_i}{\hat{p}^*}\right) + (\hat{n}_i)^2 (\hat{p}_i)^2 \left(\frac{1}{\hat{p}^*} - 1\right) \right]}{1 + c \hat{p}^*} \quad (3.3.9)$$

Several observations can be made from this expression. First, it can be seen that:

- 1) as  $c \rightarrow 0$ , we have  $\text{Var}(N_i) \rightarrow \hat{n}_i \hat{p}_i (1 - \hat{p}_i)$ , which is the variance of the best-fitting Binomial distribution to the data from line  $i$ .
- 2)  $\text{Var}(N_i)$  is an increasing function of  $c$ .
- 3) as  $c \rightarrow \infty$ , we have  $\text{Var}(N_i) \rightarrow \hat{n}_i \hat{p}_i \left(1 - \frac{\hat{p}_i}{\hat{p}^*}\right) + (\hat{n}_i)^2 (\hat{p}_i)^2 \left(\frac{1}{\hat{p}^*} - 1\right)$ .

At first glance, observation 3) may not seem very informative. However, a little effort reveals that the variance behaves as required. Specifically, Appendix 9.10 shows that, when  $c = 0$ ,  $\text{Var}(N_i)$  equals that of the best-fitting Binomial, and further,  $\text{Var}(N_i)$  is an increasing function of  $c$ . Figure 3, below, shows the behavior of the Binomial contagion variance, as a function of  $c$ . The graphs also depict cases when  $\hat{p}_i < \hat{p}^*$  as well as  $\hat{p}_i = \hat{p}^*$ , for several values of  $\hat{p}^*$ . All graphs use  $\hat{n} = 5$ .

Figure 3: Variance vs.  $c$ , per level of Binomial parameters, and  $p^*$ .



Next we investigate the correlation between  $N_i$  and  $N_j$  under the Binomial contagion model. In Appendix 9.11 it is shown that the covariance between the number of claims in the  $i^{th}$  and  $j^{th}$  lines of business, for  $1 \leq i, j \leq K$ , and  $i \neq j$ , is:

$$Cov(N_i, N_j) = \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \left( \frac{c(1 - \hat{p}^*)}{1 + c\hat{p}^*} \right) \quad (3.3.10)$$

Hence, the correlation is:

$$\rho_{N_1, N_2} = \frac{\hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \left( \frac{c(1 - \hat{p}^*)}{1 + c\hat{p}^*} \right)}{\sigma_i \sigma_j} \quad (3.3.11)$$

Where we have, from Equation (3.3.9):

$$\sigma_i = \sqrt{\frac{\hat{n}_i \hat{p}_i (1 - \hat{p}_i) + c\hat{p}^* \left[ \hat{n}_i \hat{p}_i \left( 1 - \frac{\hat{p}_i}{\hat{p}^*} \right) + (\hat{n}_i)^2 (\hat{p}_i)^2 \left( \frac{1}{\hat{p}^*} - 1 \right) \right]}{1 + c\hat{p}^*}}.$$

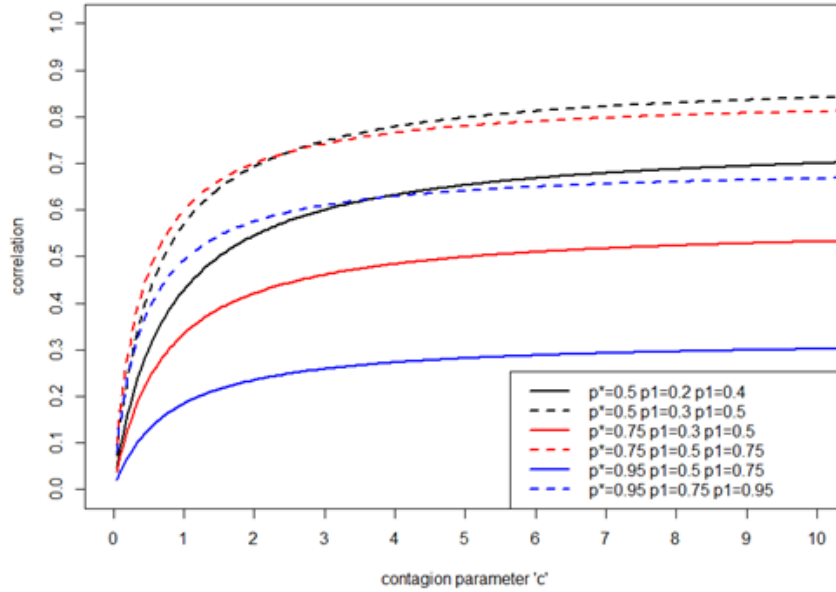
From these formulas several observations regarding  $\rho_{N_i, N_j}$ , as a function of  $c$ , can be made. For  $1 \leq i, j \leq K$  and  $i \neq j$ , we have:

- 1) as  $c \rightarrow 0$ , we have  $\rho_{N_1, N_2} \rightarrow 0$ .
- 2)  $\rho_{N_1, N_2}$  in an increasing function of  $c$ .
- 3) as  $c \rightarrow \infty$ , Appendix 9.12 shows that:

$$\rho_{N_i, N_j} \rightarrow \frac{1}{\sqrt{\left(1 + \frac{(\hat{p}^* - \hat{p}_i)}{\hat{n}_i \hat{p}_i (1 - \hat{p}^*)}\right) \left(1 + \frac{(\hat{p}^* - \hat{p}_j)}{\hat{n}_j \hat{p}_j (1 - \hat{p}^*)}\right)}} \quad (3.3.12)$$

Figure 4, below, illustrates the pair-wise correlation, under Binomial contagion, for various levels of  $\hat{p}^*$ ,  $\hat{p}_1$ , and  $\hat{p}_2$ .  $\hat{n} = 5$  is used throughout:

**Figure 4: Correlation vs.  $c$ , per level of Binomial parameters, and  $p^*$ .**



As for frequency contagion within the Poisson and Negative Binomial distributions, a single contagion parameter is shared by all Binomial claim count RV's that are subject to the same contagious environment. However this single contagion parameter will induces different levels of correlation between pairs of RV's based on the parameter values of the best fitting Binomial distributions  $\hat{n}_i, \hat{p}_i$ , and  $\hat{n}_j, \hat{p}_j$  for each line. From Figure 4, it can be seen that the correlation is higher between lines which both have relatively large  $\hat{p}_i$  and  $\hat{p}_j$ . This makes sense, since for a



given  $n$  ( $\hat{n} = 5$  in this case), larger  $p$  implies a larger number of counts from each line, and as argued above, two lines with a high number of claim counts are more likely to be highly correlated than two lines with lower counts. Also note that, the pairs of black, red, and blue lines, represent the cases of  $\hat{p}^* = 0.5$ ,  $\hat{p}^* = 0.75$ , and  $\hat{p}^* = 0.95$ , respectively. Hence, for the particular levels of  $\hat{p}_j$  in this example, it can be seen that higher value of  $\hat{p}^*$  result in lower correlations. Equation (3.3.12), illustrates why this is the case. For, unless it is the case that  $\hat{p} = \hat{p}_1 = \dots = \hat{p}_K = \hat{p}^*$ , at least one difference  $\hat{p}^* - \hat{p}_i$ , in Equation (3.3.12) will be non-zero, and as  $\hat{p}^*$  approaches 1, the corresponding term  $(\hat{p}^* - \hat{p}_i)/\hat{n}_i\hat{p}_i(1 - \hat{p}^*)$  will go to infinity. Further, since the denominator of Equation (3.3.12) consists of the product of terms of the following form;  $1 + (\hat{p}^* - \hat{p}_j)/\hat{n}_j\hat{p}_j(1 - \hat{p}^*)$ , as  $\hat{p}^*$  approaches 1, the denominator goes to infinity, while the numerator is constant. Hence, if it is *not* the case that  $\hat{p} = \hat{p}_1 = \dots = \hat{p}_K = \hat{p}^*$ , then Equation (3.3.12) goes to infinity, as  $\hat{p}^*$  approaches 1.

## 4 Severity Contagion

In Section 3 we saw that incorporating contagion within frequency models facilitates the modeling of correlation between claim count distributions. In this section we discuss the inclusion of contagion within severity distributions. It will be shown that incorporating contagion within several severity distributions not only allows the modeling of correlation between severity distributions, but also accounts for a within-line phenomenon which Actuaries and other professionals in the insurance industry have noticed for some time. Specifically, it has been observed that insurance losses from a single line often display a kind of *pseudo-correlation*, in that losses within a given period of time often tend to be uniformly higher or lower. Moreover, there are practical implementation issues which make the use of industry standard methods, such as Copula, prohibitive for modeling this phenomenon.

In traditional severity modeling, the assumption that all severities are independent is largely made to make both the mathematics and implementation easier. This is done despite the fact that practitioners are aware that losses often exhibit a pattern over time. On an industry level, the pattern of periodic increase and subsequent decline in premium rates for the same risk exposure is well known, and often referred to as the *underwriting cycle*. Moreover, it is widely believed that loss trends are a significant, if not the most significant, driver of premium rates. In turn, loss trends are driven by a set of endogenous factors such as changes in underwriting standards and policy terms, and exogenous factors which include the legal and regulatory environment. In the severity contagion modeling framework, all drivers of loss trend are subsumed within a meta-environment - the *contagious environment* – and accounted for by a single parameter, the contagion parameter. As such, a great advantage of the contagion modeling framework is the ability to describe a complex relationship between many variables with a single contagion parameter. As with frequency contagion, there is likely more than one way to incorporate contagion within severity distributions. In this paper we propose a method

similar to the method in (Meyers, 1997), insofar as the loss size RV has a multiplicative relationship with the contagion RV.

Let  $X_k$  represent the random variable for loss size, given a claim, from the  $k$ -th line of business, and let  $X_{i,k}$  represent the  $i$ -th loss from the  $k$ -th line of business. If using traditional collective risk modeling, we would assume that all losses within the same line of business are independently and identically distributed, and proceed to find the best-fitting distribution to the data from each line. Let  $RanDist_k$  represent the best-fitting distribution, and assume  $X_k$  follows distribution  $RanDist_k$ , i.e:

$$X_k \sim RanDist_k(E[X_k] = \mu_k, Var[X_k] = \sigma_{x_k}^2) \quad (4.1)$$

Based on the RV,  $X_k$ , the form of severity contagion currently in the literature proceeds by introducing a contagion RV,  $\beta$ , in a multiplicative fashion. More precisely, the product,  $\beta X_k$ , is used to simulate losses. As with frequency contagion, the only requirements on the RV  $\beta$  are that:

- 1) the distribution of  $\beta$  has positive support (i.e.  $\beta$  only takes on positive values),
- 2) the mean of  $\beta$  is 1, so that:

$$E[\beta X_k] = E[X_k] \quad (4.2)$$

Also, as with frequency contagion, the severity contagion parameter,  $b$ , is defined as the variance of  $\beta$ . Aside from these requirements,  $\beta$  can have any distributional form, with positive support. Again, we represent that  $\beta$  satisfies these requirements by writing:

$$\beta \sim RanDist(E(\beta) = 1, Var(\beta) = b) \quad (4.3)$$

Hence, this *preliminary version* of severity contagion can be summarized as follows: losses are modeled by  $\beta X_k$ , where:

- 1)  $X_k \sim RanDist_k(E[X_k] = \mu_k, Var[X_k] = \sigma_{x_k}^2)$ , where  $RanDist_k$  is the best-fitting distribution to the data from the  $k$ -th line of business.
- 2)  $\beta \sim RanDist(E(\beta) = 1, Var(\beta) = b)$ , where  $b$  is the *contagion parameter*.

We now proceed to the version of severity contagion that is proposed by the authors. The proposed form of severity contagion differs from the form currently in the literature, which we refer to as the *preliminary version*, by supposing that the size of each loss can be decomposed into two components. The first component is governed, solely, by a theoretical, unadulterated, loss distribution, and the second component represents the impact of the contagious environment on the size of loss. The former can be loosely thought of as an idiosyncratic component, and the latter an systematic component. We denote the idiosyncratic component by the RV,  $Z_k$ , and the systematic component by,  $\beta$ . Specifically, we assume that  $X_k$  can be

decomposed into the product of two RV;  $\beta Z_k$ , where  $\beta$  is the *severity contagion* RV, and  $Z_k$  the *underlying loss* RV.

The motivation for this decomposition of  $X_k$  arises mainly from the desire to separate the portion of the variation of  $X_k$  which is solely due to the random claims generation process from the portion of the variance that is due to the contagious environment. It should be noted that the proposed decomposition of the observed losses requires the elucidation of information from the data over, and above, that required by existing severity contagion models. As is usually the case, there is a cost associated with this additional precision. In this case, the additional cost comes in the form of increased complexity in calibrating the model. Specifically, instead of only the need to determine the best-fitting contagion parameter,  $b$ , it is also necessary to determine the RV,  $Z_k$ , which best describes the *pure loss process*.

As for the preliminary version of severity contagion, the only requirements on  $\beta$  are that its distribution has positive support, and that its mean be 1. Aside from this,  $\beta$  can have any distributional form. Also, since the observed losses are modeled by the product of RV's,  $\beta Z_k$ , additional requirements must be imposed, namely that:

- 1)  $\beta$  be *independent* of  $Z_k$ , and
- 2)  $E(Z_k) = E(X_k)$

In summary, the proposed severity contagion method models the size of observed losses by:

$$\beta Z_k, \text{ where: } \begin{cases} Z_k \sim \text{RandDist}_k(E[Z_k] = \mu_k, \text{Var}[Z_k] = \sigma_{z_k}^2) \\ \beta \sim \text{RandDist}(E(\beta) = 1, \text{Var}(\beta) = b), \text{ with } b \geq 0 \end{cases} \quad (4.4)$$

and, such that;  $E(Z_k) = \mu_{z_k} = \mu_k = E(X_k)$  and  $\text{Var}(Z_k) = \sigma_{z_k}^2 \leq \sigma_{x_k}^2 = \text{Var}(X_k)$ . Further, for the purposes of modeling, it is usually sufficient to assume that  $Z_k$  is a member of the same family of distributions as  $X_k$ . It can be checked that proceeding model formulation, in Equation (4.4), has the desired characteristics. First, we show that the mean of  $X_k$  is preserved by the proposed model, since:

$$E[\beta Z_k] = E[\beta]E[Z_k] = 1 \cdot \mu_{z_k} = \mu_k \quad (4.5)$$

Also, it is shown in the Appendix 9.13 that the variance has the following form:

$$\text{Var}[\beta Z_k] = \sigma_{z_k}^2 + b(\mu_k^2 + \sigma_{z_k}^2) \quad (4.6)$$

We now consider the calibration of the proposed severity contagion method. If:

- 1) the severity model, Equation (4.4), accurately represents the impact of a contagious environment on the pure loss generation process, and;
- 2) a calibration scheme is able to effectively isolate the contribution of  $\beta$  to the variation of the data,

then, at least theoretically, it should be possible to determine parameters,  $\mu_k$ ,  $\sigma_{z_k}^2$ , and  $b$ , such that:

$$Var(X_k) \approx Var[\beta Z_k] \quad (4.7)$$

This makes sense, since:

- 1) In the absence of a contagious environment, it should be inferred that;  $\sigma_{z_k}^2 \approx \sigma_{x_k}^2$ , and hence;  $\sigma_{x_k}^2 = Var(X_k) \approx Var[\beta Z_k] = \sigma_{z_k}^2 + b(\mu_k^2 + \sigma_{z_k}^2)$ , which implies that;  $b \approx 0$ , or that  $Var[\beta]$  is small.
- 2) Alternatively, in the presence of a strong contagious environment, it should be inferred that;  $\sigma_{z_k}^2 \ll \sigma_{x_k}^2$ , which, by the same argument, implies that  $Var[\beta] \gg 0$ , or in other words that;  $b \gg 0$ .

Conversely, under the same assumption that:  $Var(X_k) \approx Var[\beta Z_k]$ , we have that:

- 1)  $b \approx 0$  implies that  $\sigma_{z_k}^2 = Var[\beta Z_k] \approx Var(X_k) = \sigma_{x_k}^2$ , which implies a weak contagious environment, and:
- 2)  $b \gg 0$  implies that  $\sigma_{x_k}^2 = Var(X_k) \approx Var[\beta Z_k] = \sigma_{z_k}^2 + b(\mu_k^2 + \sigma_{z_k}^2) \gg \sigma_{z_k}^2$ , which implies a strong contagious environment.

It should also be pointed out that the ability to determine parameters,  $\mu_k$ ,  $\sigma_{z_k}^2$ , and  $b$ , such that:  $Var(X_k) \approx Var[\beta Z_k]$ , can be viewed as a refinement, or improvement, of the preliminary severity contagion model. This is so, since the preliminary severity contagion model, in the literature, takes the form,  $\beta X_k$ , and moreover, the distribution of  $X_k$  is ideally calibrated so that it captures *all* the variation in the empirical data. Hence, by fitting  $X_k$  to the data, and then multiplying by  $\beta$ , the preliminary severity contagion model,  $\beta X_k$ , will over-estimate the variance of the observed, per-claim, loss size data, i.e.:

$$Var[\beta X_k] = \sigma_{x_k}^2 + b(\mu_k^2 + \sigma_{x_k}^2) > \sigma_{x_k}^2 \approx Var[data] \quad (4.8)$$

However, this distortion of the marginal, per-claim, severity distributions would likely not be an issue when the severity contagion model is used solely for aggregate loss modeling, or when the practitioner believes that the observed data does not accurately represent the true, theoretical, variance.

So far, we have concentrated on the impact of severity contagion within a single business-line, or loss distribution. However, as with frequency contagion, a major motivation for the use of contagion modeling is its ability to model the relationship between two, or more, loss RV's. Let  $Dist_k$  denote the best-fitting distribution to the data from the  $k^{th}$  line of business, and  $Dist_j$  denote the best-fitting distribution to the data from the  $j^{th}$  line of business, and assume that:

$$X_k \sim \text{Dist}_k(E[X_k] = \mu_k, \text{Var}[X_k] = \sigma_{x_k}^2), \text{ and } X_j \sim \text{Dist}_j(E[X_j] = \mu_j, \text{Var}[X_j] = \sigma_{x_j}^2) \quad (4.9)$$

Further, assume that the parameters,  $\sigma_{z_k}^2$ ,  $\sigma_{z_j}^2$ , and  $b$ , can be determined such that:  $X_k \approx \beta Z_k$  and  $X_j \approx \beta Z_j$ , where  $\beta$  is independent of  $Z_k$  and  $Z_j$ , and  $Z_k$  is independent of  $Z_j$ . Further, assume:

$$\begin{cases} Z_k \sim \text{Dist}_{Z_k}(E[Z_k] = \mu_k, \text{Var}[Z_k] = \sigma_{z_k}^2) \\ Z_j \sim \text{Dist}_{Z_j}(E[Z_j] = \mu_j, \text{Var}[Z_j] = \sigma_{z_j}^2) \\ \beta \sim \text{Randist}(E(\beta) = 1, \text{Var}(\beta) = b) \end{cases} \quad (4.10)$$

Then, it can be seen that the covariance between  $\beta Z_k$ , and  $\beta Z_j$ , is (see Appendix 9.14):

$$\text{Cov}(\beta Z_k, \beta Z_j) = \mu_k \mu_j b \quad (4.11)$$

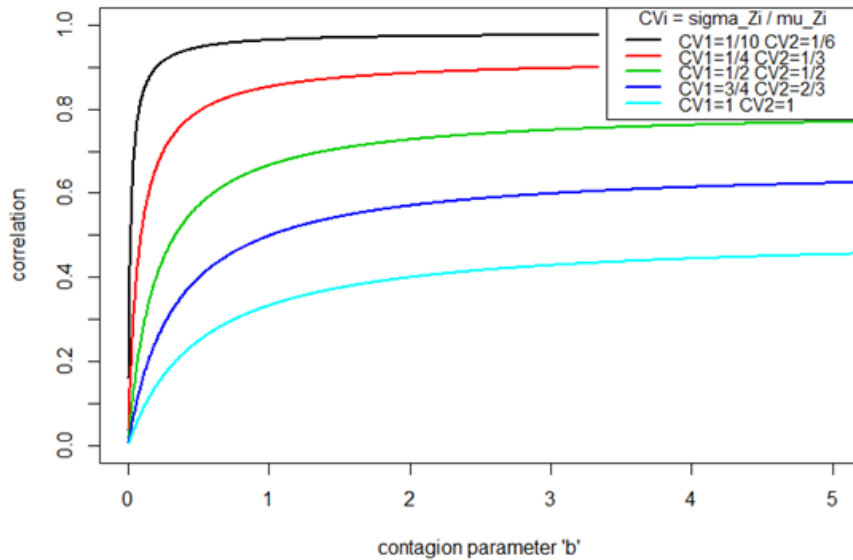
Next, in Appendix 9.15, this result is used to show that the correlation between  $\beta Z_k$ , and  $\beta Z_j$ , is:

$$\rho_{\beta Z_k, \beta Z_j} = \frac{\text{Cov}(\beta Z_k, \beta Z_j)}{\sqrt{\text{Var}(\beta Z_k) \text{Var}(\beta Z_j)}} = \frac{b}{\sqrt{[b + (1+b)CV_{z_k}][b + (1+b)CV_{z_j}]}} \quad (4.12)$$

Figure 5, below, illustrates the correlation induced between two lines of business as a function of the contagion parameter,  $b$ , and the CV's of  $Z_k$  and  $Z_j$ :  $CV_{z_k} = \sigma_{z_k}^2 / \mu_k^2$ , and  $CV_{z_j} = \sigma_{z_j}^2 / \mu_j^2$ .

From Figure 5, it can be seen that when the variation in loss size is small, the correlation which can be achieved is higher, which is consistent with intuition, and also with Equation (4.12).

**Figure 5: Correlation vs.  $b$ , and per the CV's of lines  $j$  and  $k$ .**



## 5 Aggregate loss Contagion Model

In this section we describe how the frequency contagion and severity contagion frameworks, which were developed in Sections 3 and 4, respectively, can be combined within the collective risk model. In the case studies that follow, we show that doing so results in more accurate models for aggregate losses. We first present the aggregate loss contagion model for a single line of business, and then discuss the introduction of contagion across several lines of business.

Let  $N$  be the RV which models the number of claims arising from a single line of business, within a given period of time, and let the size of each loss be modeled by the RV  $X$ . Further, let  $X_i$  denote the size of loss corresponding to the  $i^{th}$  claim, where the  $X_i$  are assumed to be *i.i.d.* Then, the model for the aggregate losses under the collective risk model is:

$$S = \sum_{i=1}^N X_i \quad (5.1)$$

Incorporating frequency and severity contagion within the collective risk model for a single line of business is straight-forward. Once the best-fitting claim count, and loss size, distributions are determined, one can simply replace  $N$ , and  $X_i$ , with the *contaged* version of  $N$ , and  $X_i$ . For the sake of argument, assume that the best-fitting frequency distribution is a Poisson with mean  $\lambda$ , such that  $N \sim \text{Poisson}(\lambda)$ , then the Poisson contagion model is:

$$N^* | C \sim \text{Poisson}(C \cdot \lambda) \text{ with } C \sim \text{Randist}(E[C] = 1, \text{Var}[C] = c) \quad (5.2)$$

Further, assume that the best-fitting severity distribution to the loss size data is,  $\text{Dist}_X(\mu, \sigma_X^2)$ , where  $\mu$ , and  $\sigma_X^2$  represent the mean and variance of the distribution of  $X$ , and that it is possible to determine parameters,  $\sigma_Z^2$ , and  $b$ , such that  $X \approx \beta Z$ , where  $\beta$  is independent of  $Z_i$ , and:

$$\begin{cases} Z_i \sim \text{Dist}_Z(E[Z_i] = \mu, \text{Var}[Z_i] = \sigma_Z^2) \\ \beta \sim \text{Randist}(E(\beta) = 1, \text{Var}(\beta) = b) \end{cases} \quad (5.3)$$

Then, we can simply replace  $N$  with  $N^*$ , and  $X_i$  with  $\beta Z_i$  within  $S$ , to arrive at the aggregate loss distribution under frequency and severity contagion, which we denote by  $S^*$ :

$$S^* = \sum_{i=1}^{N^*} \beta Z_i \quad (5.4)$$

It is shown in Appendix 9.16 that the expected value of  $S^*$  equals the expected value of the traditional collective risk model, i.e.:

$$E[S^*] = \lambda \cdot \mu \quad (5.5)$$

Also, in Appendix 9.17, it is shown that the variance of the aggregate loss, under both frequency and severity contagion, is:

$$Var[S^*] = Var[Z_i]E(N^*) + (E[Z_i])^2Var(N^*) + b \cdot \{Var[Z_i]E(N^*) + (E[Z_i])^2E(N^{*2})\} \quad (5.6)$$

Note that Equation (5.6) incorporates both frequency and severity contagion, and does not depend on the distributional form of the claim count RV (Poisson, Negative Binomial, or Binomial). One can simply substitute in the appropriate formulas for  $E(N^*)$ , and  $Var(N^*)$ , corresponding to the frequency contagion model employed. The same observation can be made regarding the severity RV  $Z$ . Further, the Equation (5.6) does not depend on the distributional form of the contagion RV's;  $C$ , or  $\beta$ . Hence, Equation (5.6) is a fully general formula.

We can now investigate the correlation between the aggregate losses of two lines,  $S_k^*$  and  $S_j^*$ , for  $k \neq j$ . In Appendix 9.18 it is shown that the covariance between the aggregate losses of two lines of business, under both frequency and severity contagion, is:

$$Cov(S_k^*, S_j^*) = E[Z_{ki}]E[Z_{ji}] \cdot \{b \cdot E(N_k^* N_j^*) + Cov(N_k^*, N_j^*)\}, \quad \text{for } k \neq j \quad (5.7)$$

In order to make the formula for the correlation between  $S_k^*$  and  $S_j^*$  more presentable, we first note that we can re-write Equation (5.6) for the  $Var[S^*]$  as:

$$Var[S^*] = Var[Z_i]E(N^*) + (E[Z_i])^2Var(N^*) + b \cdot \{Var[Z_i]E(N^*) + (E[Z_i])^2Var(N^*) + (E[Z_i] \cdot E(N^*))^2\}. \quad (5.8)$$

Or:

$$Var[S^*] = \Sigma + b \cdot [\Sigma + (\mu\lambda)^2] \quad (5.9)$$

Where:

$$\begin{cases} \Sigma = Var[Z_i]E(N^*) + (E[Z_i])^2Var(N^*) \\ \mu = E[Z_i] \\ \lambda = E(N) = E(N^*) \end{cases} \quad (5.10)$$

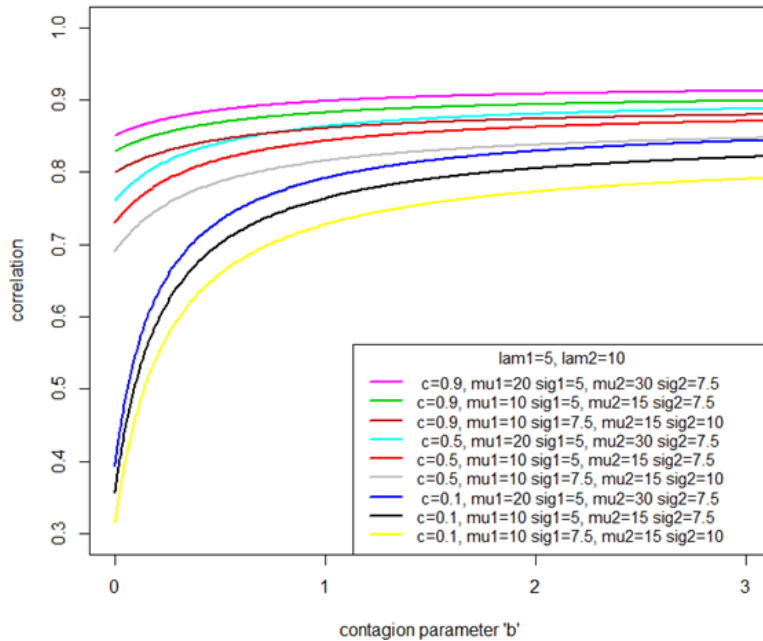
Using this notation, in Appendix 9.19, it is shown that the correlation between  $S_k^*$  and  $S_j^*$ , for  $k \neq j$ , is:

$$\rho_{S_k^*, S_j^*} = \frac{\lambda_k \mu_k \lambda_j \mu_j \cdot (cb + b + c)}{\sqrt{(\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]) (\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2])}} \quad (5.11)$$

Figure 6, below, illustrates the correlation induced between the aggregate losses of two lines of business, as a function of the contagion parameter,  $b$ . Each of the nine graphs uses 5 and 10, for the means of the frequency distributions, i.e.:  $\lambda_1 = 5$ , and  $\lambda_2 = 10$ . However, the graphs correspond to three levels of the frequency contagion parameter;  $c = 0.1, 0.5$ , and  $0.9$ . The base case, for each of these three levels of  $c$ , uses  $\mu_1 = 10, \sigma_1 = 5, \mu_2 = 15$  and  $\sigma_2 = 7.5$ . Then, for each of the three levels of  $c$ , two more plots are made, one with larger means;  $\mu_1 = 20, \sigma_1 = 5, \mu_2 = 30$  and  $\sigma_2 = 7.5$ , and one with larger variance;  $\mu_1 = 10, \sigma_1 = 7.5, \mu_2 = 15$  and  $\sigma_2 = 10$ .

It can be seen that higher levels of the frequency and severity contagion parameters;  $c$  and  $b$ , result in higher correlation between the lines. Also, higher levels of  $\mu_j$  result in higher correlation. Conversely, higher levels of  $\sigma_j$  result in *lower* correlation, however. These observations are consistent with intuition. As with severity correlation, for the same level of  $\sigma_j$ , higher levels of  $\mu_j$  result in smaller  $CV_j$ , which results in higher correlation. Also, for the same level of  $\mu_j$ , the greater the uncertainty in the loss sizes of each line, represented by  $\sigma_j$ , the more tenuous the relationship between the loss sizes, and hence the lower the correlation.

**Figure 6: Correlation of aggregate losses vs.  $b$ , per level of  $c$ , and the mean & SD of the severity distributions.**



Moreover, in the special case where the claim frequency for both line  $j$  and  $k$ , are Poisson, it is shown in Appendix 9.20 that the correlation between  $S_k^*$  and  $S_j^*$ , for  $k \neq j$ , is:



$$\rho_{S_k^*, S_j^*} = \frac{(cb + b + c)}{\sqrt{[(1 + b)CV_{S_k}^2 + [bc + b + c]] [(1 + b)CV_{S_j}^2 + [bc + b + c]]}} \quad (5.12)$$

Note that  $CV_{S_k}^2$ , in Equation (5.12), is the  $CV$  of the aggregate losses under the *traditional collective risk model*, i.e. *without* either frequency or severity contagion. However, it should be pointed-out that  $CV_{S_k}^2$  uses the  $Var(Z_k)$ , rather than  $Var(X_k) = Var[data_j]$ . Further, recall that  $Z_k$  is defined such that  $X_k \approx \beta Z_k$ , More explicitly,  $CV_{S_j}^2$  and  $CV_{S_k}^2$  are the  $CV$ 's of the following  $RV$ 's:

$$S_j = \sum_{i=1}^{N_j} Z_{j_i}, \quad \text{and} \quad S_k = \sum_{i=1}^{N_k} Z_{k_i}$$

where:

$$\left\{ \begin{array}{l} N_j \sim \text{Poisson}(\lambda_j) \\ Z_{j_i} \sim \text{Dist}_{Z_j}(\mu_j, \sigma_{Z_j}^2) \end{array} \right\}, \quad \text{and} \quad \left\{ \begin{array}{l} N_k \sim \text{Poisson}(\lambda_k) \\ Z_{k_i} \sim \text{Dist}_{Z_k}(\mu_k, \sigma_{Z_k}^2) \end{array} \right\}$$

Where:  $Var(Z_j) = \sigma_{Z_j}^2 \leq \sigma_{X_j}^2 = Var(X_j) = Var[data_j]$ .

## 6 Implementation & Calibration:

In this section we describe a method for calibrating contagion models for aggregate losses, using empirical data. At the time of this writing, there is a dearth of literature on this topic, both within the peer-reviewed literature, and publically available insurance industry literature. To the authors knowledge, this paper represents the first attempt to document, in detail, calibration procedures for contagion models based on empirical, *per-claim*, loss data. In addition, a new severity contagion calibration method is proposed. The proposed severity calibration procedure is described for a single line of business. Armed with an understanding of the proposed calibration method for a single line of business, various calibration procedures can be employed to extend the calibration to two, or more, lines of business. We leave the choice of the specific calibration procedure to the practitioner. The procedure employed will, no doubt, depend on the purpose of the analysis as well as the expert opinion of the practitioner.

There are three main steps for calibration of contagion models for aggregate losses:

1. Model fitting - determination of the best-fitting distributions for the frequency and severity, based on the empirical data.
2. Calibration of the frequency contagion model, based on the empirical claim count data.
3. Decomposition of the variance of the severity distribution, resulting in calibrated models for both the pure, underlying, loss process,  $Z_i$ , and also the contagion  $RV \beta$ .

Steps 1 is not distinct from traditional frequency and severity modeling, and can be completed using MLE, method-of-moments, or any other estimation procedure. As should always be the case, the characteristics of estimates produced by the chosen method should be kept in mind. In particular, it should be kept in mind that in many situations MLE estimates are biased. If it is determined that producing unbiased estimates of aggregate losses is of up-most importance, then it may be desired to use method-of-moments for calibration of the frequency model.

Step 2 is accomplished by exploiting the corresponding formula for the variance of the claim count RV under contagion. For example, when claim counts are modeled with a Poisson distribution, with mean  $\lambda$ , we have from Equation (3.1.7) that:  $Var(N) = \lambda(1 + c \cdot \lambda)$ . As  $Var(N)$  can be estimated from the observed annual claim count data, the only unknown quantity in Equation (3.1.7) is the frequency contagion parameter  $c$ . Hence, Equation (3.1.7) can be re-written to solve for  $c$ :

$$c = \frac{Var(N)}{\lambda^2} - \frac{1}{\lambda} \quad (6.1)$$

Step 3, is unique to the proposed severity calibration method, and requires additional care. As explained in Section 4, the proposed severity calibration requires the identification of a RV  $Z_k$ , such that the total variation of  $X_k$  can be sufficiently captured by the product of the contagion RV  $\beta$ , and the underlying loss RV,  $Z_k$ . Hence, the proposed severity calibration method requires the elucidation of more information from the data than is required under the severity contagion calibration procedures which exist in the literature. However, recall that we assume the RV  $Z_k$  follows the same distributional form as the best-fitting distribution to  $X_k$ , and hence the determination of the distribution of  $Z_k$  only requires the determination of its variance,  $\sigma_z^2$ . If sufficient data are available, one way to accomplish this would be to use advanced statistical procedures, such as Analysis of Variance (ANOVA), GLMs, or Factor Models. Though the use of such techniques may be the more statistically correct way to proceed, documentation of the application of these procedures for this type of calibration is worthy of a separate treatment, which the authors hope to complete in the near future. Regardless, it is possible to perform the proposed calibration without needing to rely on such powerful statistical techniques. Such a straight-forward calibration procedure also has the added benefit of being more palatable to industry practitioners. Also, note that the proposed calibration method is informed by the goal of the case studies in Section 7; namely the investigation of the variation in aggregate annual layered losses. We now describe the proposed severity calibration method.

To accomplish Step 3 we perform the following steps:

- 1) Determine  $b$
- 2) Use this  $b$  to determine  $\sigma_z^2$
- 3) Determine the parameters of the severity distribution (Pareto, in this case), which correspond to this  $\sigma_z^2$ , and use this severity distribution within the simulation.

From Section 5 we know that the variance of the aggregate annual losses, under frequency and severity contagion, is:

$$Var[S^*] = Var[Z_i]E(N^*) + (E[Z_i])^2Var(N^*) + b \cdot \{Var[Z_i]E(N^*) + (E[Z_i])^2E(N^{*2})\} \quad (6.2)$$

If claim frequency is assumed to follow a Poisson distribution with parameter  $\lambda$ , then:

$$\begin{aligned} Var[S^*] &= \sigma_z^2 \lambda + \mu_z^2 \lambda(1 + c\lambda) + b \cdot \{\sigma_z^2 \lambda + \mu_z^2 [\lambda(1 + c\lambda) + \lambda^2]\} = \\ &= \lambda(1 + b)(\mu_z^2 + \sigma_z^2) + \lambda^2 \mu_z^2 [b + c + bc]. \\ \rightarrow b &= \frac{Var[S^*] - \lambda \sigma_z^2 - \lambda b(\mu_z^2 + \sigma_z^2) - \lambda \mu_z^2 - \lambda^2 \mu_z^2 c}{\lambda^2 \mu_z^2 (1 + c)} \end{aligned} \quad (6.3)$$

Also, from Equation (4.6), we **know** have that:

$$Var[\beta Z_k] = \sigma_z^2 + b(\mu_z^2 + \sigma_z^2) \quad (6.4)$$

Hence we have that:

$$b = \frac{Var[S^*] - \lambda \cdot Var[\beta Z] - \lambda \mu_z^2 - \lambda^2 \mu_z^2 c}{\lambda^2 \mu_z^2 (1 + c)} \quad (6.5)$$

In Equation (6.5),  $Var[S^*]$  is the variance of the empirical aggregate losses, and  $\mu_k$  and  $Var[\beta Z] \approx Var[X]$  can be estimated using the mean and the variation, respectively, of the *per-claim* loss data. Recalling that  $c$  has already been estimated, at this stage, all remaining quantities to the right of the equals sign, in Equation (6.5), can be estimated from the empirical data. Note that Equation (6.5) only applies to the specific case of a Poisson frequency. However, when contagion is introduced within a Negative Binomial frequency distribution  $b$  can be still be derived from Equation (6.2), by instead using the applicable formulas for  $Var(N^*)$  and  $E(N^{*2})$  within Equation (6.2). In this case, the formula for  $b$  can be shown to be:

$$b = \frac{Var[S^*] - \lambda \cdot Var[\beta Z] - \lambda \mu_z^2 - \lambda^2 \mu_z^2 c - \lambda^2 \mu_z^2 \gamma(1 + c)}{\lambda^2 \mu_z^2 (1 + c + \gamma + c\gamma)} \quad (6.6)$$

Once  $b$  has been determined, the variance of the pure, underlying, loss  $RV$ ,  $Z$ , can be solved for using Equation (6.4) together with the value of  $b$ :

$$\sigma_z^2 = \frac{Var[\beta Z] - b\mu_z^2}{1 + b} \approx \frac{Var[X] - b\mu_z^2}{1 + b} \quad (6.7)$$

Finally, with values for:

- the parameters of the frequency distribution,
- the frequency contagion parameter  $c$ ,
- the severity contagion parameter  $b$ , and
- the mean and variance of  $RV$ ,  $Z$ ;  $\mu_z$  and  $\sigma_z^2$

the aggregate annual losses, under both frequency and severity contagion, can be simulated using:

$$S^* = \sum_{i=1}^{N^*} \beta Z_i \quad (6.8)$$

where:

$$N^* | C \sim \text{Dist}_N(C \cdot \lambda) \text{ with } C \sim \text{Randist}(E[C] = 1, \text{Var}[C] = c) \quad (6.9)$$

and:

$$\begin{cases} Z_i \sim \text{Dist}_X(E[Z_i] = \mu_z = \mu_x, \text{Var}[Z_i] = \sigma_z^2) \\ \beta \sim \text{Randist}(E(\beta) = 1, \text{Var}(\beta) = b) \end{cases} \quad (6.10)$$

As indicated above, this is just one method for calibrating contagion models. This particular method is informed by the desire to produce aggregate annual loss simulations with variation close to the variation of the empirical annual aggregate losses. Another advantage of this procedure is that it is relatively straight-forward, and easy to implement. In Section 7 case studies, which investigate the variation in simulated aggregate annual layered losses based on traditional collective risk models verses those based on contagion models, are presented.

## 7 Case Study

In this section, case studies illustrating the application of contagion to aggregate layered loss modeling been performed on two sets of data. Both case studies utilize actual industry data, and moreover, both case studies illustrate the calibration of the proposed contagion models to historical industry data. The first data set are property natural peril losses from a US insurer, and the second data set is from a US general liability (GL) line of business. Each data set corresponds to a different company. As the data is proprietary, the company names are withheld.

The case studies below illustrate the application of frequency and severity contagion for aggregate loss modeling, within a single line of business. For each data set, the calibration procedure described in Section 6 are followed. The analysis begins with a cursory vetting of both data sets for reasonableness, and to make sure that all parameters of the data, including any potential adjustments, such as trending, are understood. Once the data are fully understood, the claim counts are aggregated, by year. Using this annual claim count data, the best-fitting family of frequency distributions is determined. This can be accomplished by using goodness-of-fit statistics, such as *AIC*, or using Bayesian techniques, such as Bayes Factors. Once the distributional form of the annual claim count data is determined, the parameter values are determined from sample estimates, equivalent to using the method-of-moments (MoM). Method-of-moments estimates are used, instead of MLE estimates in the final parameterization, out of the desire to generate simulations which result in estimates of the first and second moment of aggregate losses that match the empirical estimates.

Next, the contagion parameter  $c$  for the frequency contagion model is determined. Again, the authors are unaware of exist literature which describes, in detail, how to perform this calibration using per-claim data. As is the case with any form of modeling, the calibration procedure for contagion models depends, in part, on the purpose of the study. Both of the case studies presented compare the variation in the simulated aggregate annual layered losses under the traditional collective risk model with the same quantities under the contagion model. In both case studies, the calibration of the frequency contagion model reflects this fact. Also, in both case studies, the severity decomposition method, proposed by the authors, is implemented. As a reminder, this means that the observed losses are modeled as the product of two RVs;  $Z_k$  and  $\beta$ , whose distributions must both be calibrated. Using statistical analysis, the parameter values of the theoretical, pure, underlying loss  $Z_k$  distribution, and the contagion parameter  $b$  corresponding to the contagion RV, are both educed from the data, simultaneously.

## 7.1 Case Study #1: Property Natural Peril Severity Data

The first set of data consists of 670 claims between years 2003 and 2012, inclusive. The data is on an occurrence basis, and is from a single company. Additional details regarding the data are as follows:

- Property Natural Peril (severe convective storm)
- For a single company.
- Occurrence basis.
- Losses over 10-years: 2003 – 2012
- Loss sizes are in units of \$1,000.

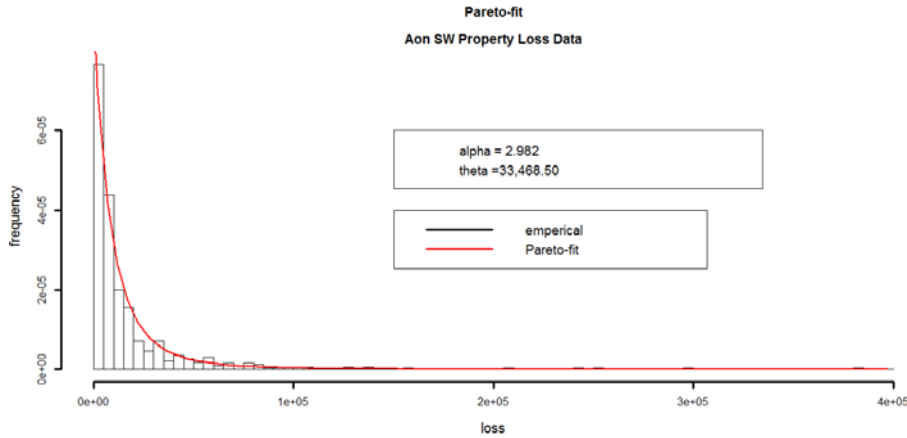
We use a Poisson distribution for the claim counts, with parameter value  $\lambda$  set equal to  $\bar{X} = 67$ . We now calibrate the contagion model by following the steps in Section 6. Using  $E(N) = \lambda = 67$  and  $Var(N) = 604$  within Equation (6.1), we determine the frequency contagion parameter:

$$c = \frac{Var(N)}{\lambda^2} - \frac{1}{\lambda} \approx 0.115$$

At this point the frequency component of the contagion model is parameterized. We now turn our attention to calibrating the severity component of the contagion model. The best-fitting distribution to the per-claim severity is a Pareto distribution with  $\alpha_{MLE} = \text{shape parameter} = 2.982$ , and  $\theta_{MLE} = \text{scale parameter} = 33,468$  (for details, see Appendix 9.21). The fit of the Pareto distribution to the empirical data is illustrated in Figure 7. It can readily be seen that this Pareto provides a good fit to the data. As mentioned in Section 6, the severity distribution based on the MLE estimates, can be used within the simulation. However, we use the Pareto distribution based on the method-of-moment estimates ( $\{\alpha_{MoM}, \theta_{MoM}\} = \{2.876, 33,470\}$ ), in order to produce unbiased estimates of the mean and variance. The chart in Table 1 provides a comparison of the severity parameterization using MLE verse method-of-moments. Note that the mean and standard deviation corresponding to the MLE estimates deviate from sample-

based estimates. This is due to the general *biasedness* of MLE estimates, when based on finite sample sizes. Next, as we remarked earlier, in both case-studies, we implement the proposed severity calibration procedure (as described in Section 6), which involves the decomposition of the observed severity variance into two components; the variance of the theoretical underlying loss  $RV$ ,  $Z$ , and the variance of the severity contagion parameter  $b$ . Further, we assume that the  $RVZ$  follows the same distributional form as the best-fitting distribution to  $X$ , which is Pareto in this case. We now calibrate the severity component of the contagion model.

**Figure 7: Comparison of fitted and empirical densities of the severity data.**



**Table 1: MLE & MoM estimates of severity distribution parameters**

Method of estimation	Pareto parameters		standard deviation	
	alpha	Theta	mean	
MLE	2.982	33,468	16,881	29,413
MoM	2.876	33,470	17,842	32,329

First, we use Equation (6.5) to determine the severity contagion parameter. The input to Equation (6.5) are:

- The mean of the Poisson frequency distribution ( $\lambda = 67$ ).
- The modeled mean of the per-claim severity distribution ( $\mu_z = 17,842$ ).
- The modeled variance of the per-claim severity ( $Var[\beta Z] = \sigma_x^2 = 32,329^2$ ).
- The *empirical* variance of the annual aggregate losses (AAL): ( $Var[S^*] = 697,245^2$ ).
- The frequency contagion parameter ( $c = 0.115$ ).

$$b = \frac{Var[S^*] - \lambda \cdot Var[\beta Z] - \lambda \mu_z^2 - \lambda^2 \mu_z^2 c}{\lambda^2 \mu_z^2 (1 + c)} \approx 0.13$$

Next, we use Equation (6.7), together with the value of  $b$ , to determine the variance of the pure, underlying, loss  $RV$ ,  $Z$ :

$$\sigma_z^2 = \frac{Var[\beta Z] - b \mu_z^2}{1 + b} = 29,634^2$$

If a statistical software is used, which does not allow the parameterization of distributions using the mean and variance, the parameter values can be solved for in terms of the mean and variance. For the Pareto distribution, the corresponding formulas are:

$$\alpha = \frac{2Var[Z]}{Var[Z] - E[Z]^2} \quad (7.1)$$

$$\theta = E[Z] \left( \frac{Var[Z] + E[Z]^2}{Var[Z] - E[Z]^2} \right) \quad (7.2)$$

The resulting parameters are:  $\alpha_z = 3.137$ , and  $\theta_z = 38,133$ . At this point the severity contagion calibration is completed, and we can begin the simulations, which are at the core of this study. First, we define the layers of loss, which will be used for all simulation in case study #1:

$$\left\{ \begin{array}{l} \text{Layer1: } 0 - 7.5M \\ \text{Layer2: } 7.5M - 20M \\ \text{Layer3: } 20M - 45M \\ \text{Layer4: } 45M - 70M \\ \text{Layer5: } 70M - 100M \\ \text{Layer6: } 100M - 200M \end{array} \right. \quad (7.3)$$

Next, for each iteration of the simulation study, we simulate 10 years of aggregate annual layered losses (AALL). Then we calculate the coefficient-of-variation ( $CV$ ) of this set of 10 aggregate annual losses. We do this both using simulated aggregate annual layered losses under contagion, and simulated without contagion. This process is repeated 100,000 times, which results in 100,000  $CV$  estimates simulated under contagion, and 100,000  $CV$  estimates simulated without contagion. Both sets of 100,000 simulated  $CV$ 's are then displayed in a box-plot, along with the empirical aggregate annual layered losses. For reference, the parameters of the distributions used to simulate the aggregate annual layered losses, both under contagion, and without contagion are displayed in Table 2. The complete procedure is as follows:

1. For each iteration of the simulation, generate 10 years of Aggregate Annual Losses under both:
  - The Traditional method
  - The Contagion method.

**Table 2: Parameters of the distributions used to simulate AAL's, both under contagion and without contagion.**

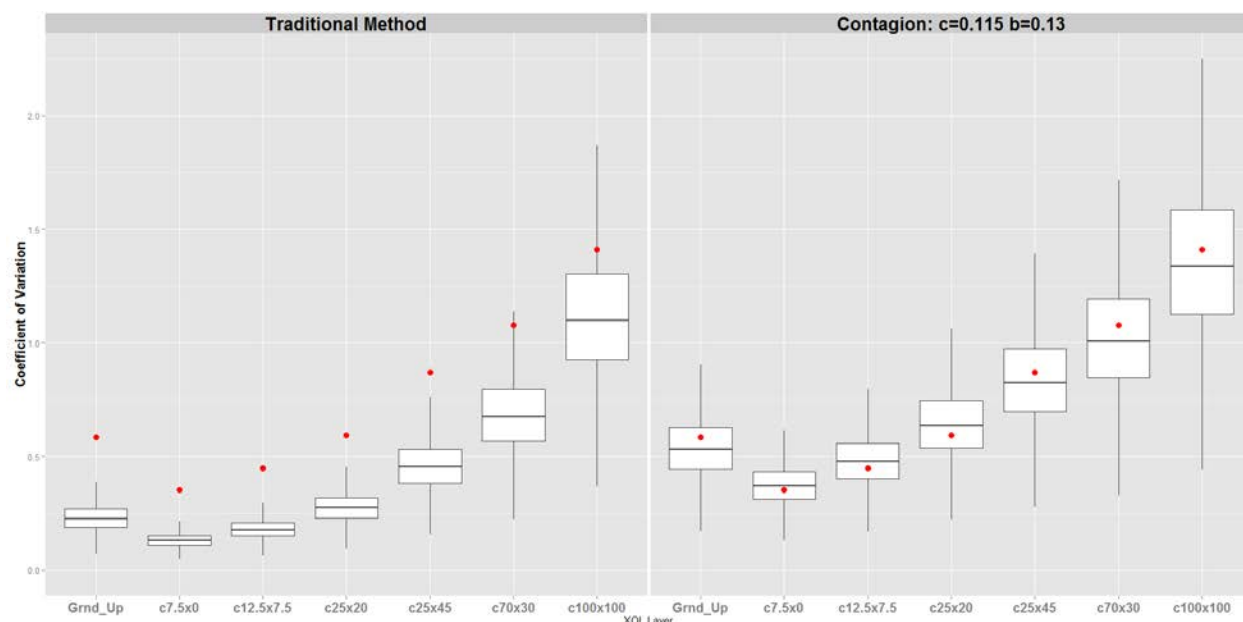
	Traditional Collective Risk Model: Poisson – frequency & Pareto - severity	Aggregate Contagion model: Poisson – frequency & Pareto - severity
Frequency mean	$\lambda = 67$	$\lambda = 67$
Severity shape parameter	$\alpha = 2.876$	$\alpha_z = 3.137$
Severity scale parameter	$\theta = 33,470$	$\theta = 38,133$
Frequency Contagion parameter	$N/A$	$c = 0.115$
Severity Contagion parameter	$N/A$	$b = 0.13$

2. For each of the pre-defined Loss Layers (Equation (7.3)), calculate the Annual Aggregate losses within each layer.
3. For each layer, calculate the  $CV$  of the Aggregate Annual losses, over the 10-years.
4. Repeat 100,000 times.
  - At this point, we have 100,000 10-year  $CV$  estimates, for each layer.
5. For each layer, order the 100,000 10-year  $CV$  s, and compute the 10th, 25th, 50th, 75th, and 90th percentiles of the set of 100,000 10-year  $CV$ 's.
6. Finally, box-plots of the simulated  $CV$ 's for each layer are plotted. The ground-up losses, are on the far left of each graph. The empirical  $CV$ 's of the aggregate annual layered losses are also computed, and shown in red, in Figure 8.

Figure 8 illustrates that using the traditional Collective Risk Model (left) results in drastically understating the  $CV$  of losses, in all layers, as well the  $CV$  of the ground up losses. On the other hand, the Aggregate Contagion model (right), calibrated using the straight-forward technique described in Section 6, produces estimates of the aggregate annual layered losses whose  $CV$  is much closer to the empirical  $CV$  of the layered losses. It is important to note that this occurs, uniformly, for all layers. Specifically, a single set of contagion parameters,  $c$  and  $b$ , can be used to more accurately model the variation of losses, within all layers, simultaneously. Within the insurance industry, the ability to accurately capture variation, instead of just the mean, has grown in recognition over the past several decades. In particular, risk measures such as *Value-at-Risk* ( $VaR$ ) and *Conditional Tail Expectation* ( $CTE$ ) or *Tail-VaR* ( $TVaR$ ), were constructed to capture such distributional qualities. The use of contagion within the traditional collective risk model will allow Actuaries, and risk professionals, to reap the benefits of these modern risk measures.



**Figure 8: Box plots of CVs of AALL's under both the Traditional Collective Risk model, and the Contagion model.**



## 7.2 Case Study #2: General Liabilities (GL) claims data:

The 2<sup>nd</sup> case study utilizes general liabilities (GL) data, details are as follows:

- LOB: GL claims
- Losses over 7-years: 2007 – 2013
- For a single company
- On a per-occurrence basis
- Losses recorded after policy-limits
- Loss amounts are in units of \$1

In case study #2, the average of the empirical annual claim counts is approximately 8,500. Further, the variance of the empirical annual claim counts is over 100 times greater than the mean. Hence, in this case study we choose to model the annual claim counts with a Negative Binomial distribution. It should be noted, however, that if the claim counts were modeled, instead, with a Poisson contagion model, as in case study #1, the frequency contagion parameter within the Poisson frequency contagion model will automatically account for the extra dispersion in empirical annual claim counts. The use of the Negative Binomial contagion model,

however, also provides an opportunity to discuss some additional, and potentially subtle, issues which arise when using the Negative Binomial under frequency contagion. The reader may note that from Equation (3.2.8), for the Negative Binomial distribution, we have;  $Var(N) = \lambda(1 + \lambda(c + c\gamma + \gamma))$ , while from Equation (3.2.7), the mean is;  $E(N) = \lambda$ , as for the Poisson. Hence, when using a Negative Binomial frequency, more than two formulas are needed to uniquely determine  $\lambda$ ,  $c$ , and  $\gamma$ . However, this does not pose a problem. For instance, the formula for a higher moment, or for some other quantity of interest, can be used in addition to  $E(N)$  and  $Var(N)$ . This will produce three equations with three unknowns,  $\lambda$ ,  $c$  and  $\gamma$ . However, if  $c$  is assumed to be zero, the dispersion parameter  $\gamma$  of the Negative Binomial is equivalent to the contagion parameter under a Poisson distribution. Hence, if  $c$  is assumed to be 0,  $E(N)$  and  $Var(N)$  can be used to solve for  $\gamma$ . This is, in fact, the procedure used in this case study. This is a valid assumption, since we are only modeling the aggregate losses from a single line of business, and hence, there is no across-line correlation. Note that if more than one line of business were being modeled, each  $\gamma_i$  could be calibrated to account for the over-dispersion in the corresponding line, and  $c$  would be calibrated to account for the correlation in claim counts, across the lines.

We now calibrate the Negative Binomial frequency contagion model by solving for  $\gamma$  so that the mean and variance of the Negative Binomial match the empirical mean and variance of the claim count data. Again, we assume  $c = 0$ , and solve for  $\gamma$  as:

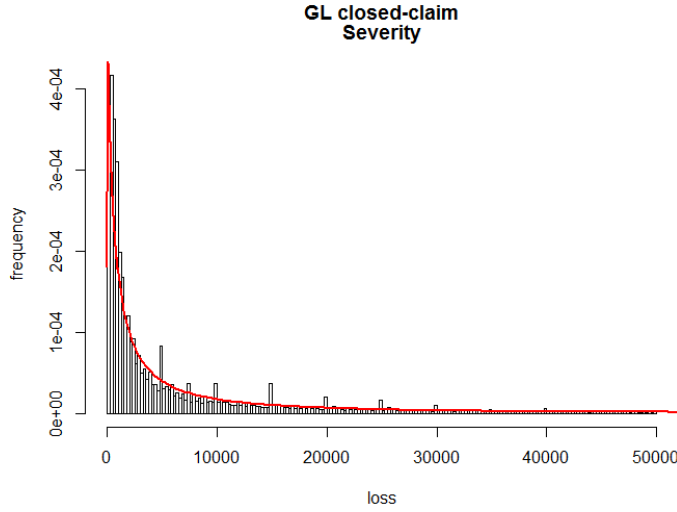
$$\gamma = \frac{Var(N)}{E(N)^2} - \frac{1}{E(N)} \approx 0.086$$

We now model the per-claim loss sizes. The LogNormal family provides the best-fit to the data in case study #2. Figure 9 shows the best fitting LogNormal distribution to the empirical density. The parameters of the LogNormal, using the method-of-moments, are  $(\mu, \sigma) = (8.65, 1.64)$ .

Next, we use Equation (6.6) to determine the severity contagion parameter. The input to Equation (6.6) are:

- The mean of the Negative Binomial frequency distribution ( $\lambda = E(N) = 8,679$ ).
- The dispersion parameter of the Negative Binomial ( $\gamma = 0.0247$ )
- The modeled mean of the per-claim severity distribution ( $\mu_z = 26,764$ ).
- The modeled variance of the per-claim severity ( $Var[\beta Z] = \sigma_x^2 = 87,657^2$ ).
- The *empirical* variance of the annual aggregate losses ( $Var[S^*] = 67,694,180^2$ ).
- The frequency contagion parameter ( $c = 0$ )

Figure 9: Comparison of fitted and empirical densities of the severity data.



Which results in:

$$b = \frac{Var[S^*] - \lambda \cdot Var[\beta Z] - \lambda \mu_z^2 - \lambda^2 \mu_z^2 \gamma}{\lambda^2 \mu_z^2 (1 + \gamma)} \approx 0.057$$

Also, as in case study #1, we use Equation (6.7), together with the value of  $b$ , to determine the variance of the pure, underlying, loss RV,  $Z$ :

$$\sigma_z^2 = \frac{Var[\beta Z] - b \mu_z^2}{1 + b} = 85,017^2$$

The corresponding LogNormal parameter values corresponding to the RV  $Z$ , are:  $(\mu_z, \sigma_z) = (9, 1.55)$ . At this point the severity contagion calibration is completed, and we can begin the simulations which are at the core of this study. First, we define the layers of loss, which will be used for all simulation, in case study #2:

$$\left\{ \begin{array}{l} \text{Layer1: } 0 - 20K \\ \text{Layer2: } 20K - 30K \\ \text{Layer3: } 30K - 60K \\ \text{Layer4: } 60K - 100K \\ \text{Layer5: } 100K - 200K \\ \text{Layer6: } 200K - 500K \end{array} \right. \quad (7.4)$$

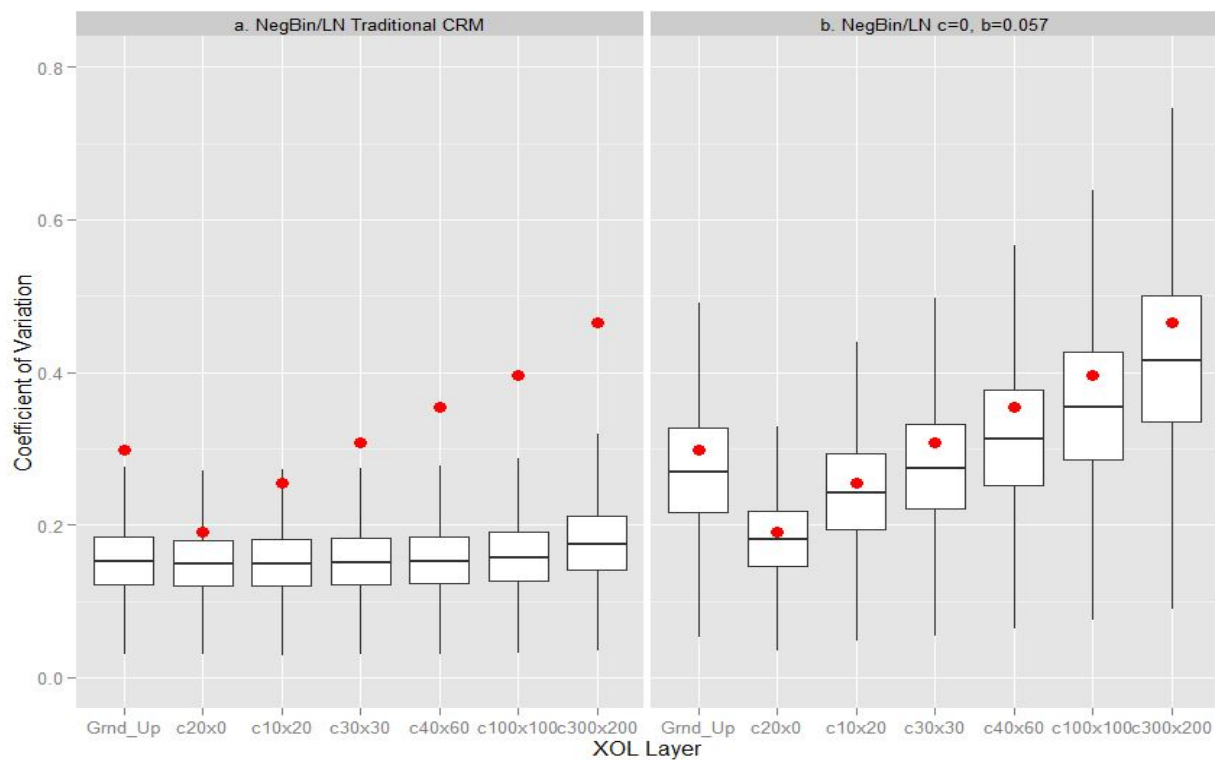
Next, just as for case study #1, for each iteration of the simulation study, we simulate 7 years of aggregate annual layered losses. Then we calculate the  $CV$  of this set of 7 aggregate annual losses. We do this both using simulated aggregate annual layered losses under contagion, and without contagion. This process is repeated 100,000 times, which results in 100,000  $CV$

estimates simulated under contagion, and 100,000  $CV$  estimates simulated without contagion. Both sets of 100,000 simulated  $CV$ 's are then displayed in a box-plot, along with the empirical aggregate annual layered losses. For reference, the parameters of the distributions used to simulate the aggregate annual layered losses, both under contagion, and without contagion are displayed in Table 3.

**Table 3: Parameters of the distributions used to simulate AAL's, both under contagion and without contagion.**

	Traditional Collective risk Model:	Aggregate Contagion model:
Frequency (Negative Binomial)	$\lambda = 8,679$	$\lambda = 8,679$
	$\gamma = 0.0247$	$\gamma = 0.0247$
		$c = 0$
Severity (Lognormal)	$\mu = 9$	$\mu = 9$
	$\sigma = 1.55$	$\sigma = 1.55$
		$b = 0.057$

**Figure 10: Box plots of CVs of AALL's under both the Traditional Collective Risk model, and the Contagion model.**



The results, shown in Figure 10, are even more dramatic than in case study #1. We can see that using the traditional collective risk model results in the  $CV$  of the aggregate annual layered losses being drastically underestimated for all layers, as well as for the ground up losses. On the other hand, the aggregate Contagion model produces estimates of the aggregate annual layered losses whose  $CV$  is much closer to the empirical  $CV$  of the layered losses. This is the case, even though a Negative Binomial distribution was used for the frequency of both the traditional collective risk model, and the Contagion model. This is significant since, as mentioned above, the Negative Binomial distribution is equivalent to using a Poisson distribution under contagion. Hence, the difference between the two sets of box-plots, in Figure 10, is solely due to the incorporation of severity contagion.

## 8 Conclusion

In this paper we have advanced the Common Shock, or Contagion, modeling paradigm, as introduced by Meyers (1997). These advancements consist of, first, the analytical elucidation of the impact of this modeling paradigm. This has been done for frequency, severity, as well as aggregate Common Shock, or Contagion, modeling. Second, we have developed the application of Contagion modeling to additional distributions. Third, we have developed calibration procedures which allow the parameterization of aggregate Contagion models based on empirical data. Last, we have demonstrated the efficacy, and ease, of the proposed calibration procedures using several, real-world, data sets from the Insurance industry. These case studies illustrate the well-known limitations of traditional collective risk modeling, and provide evidence that Contagion modeling provides a systematic way to correct for the inherent limitations of traditional collective risk modeling. It is our hope that this work will help establish Contagion modeling, both among academics and practitioners, and also motivate additional research on Contagion modeling, as well as on other ways to improve traditional collective risk modeling.

## 9 Appendix A

### 9.1 Poisson Frequency Contagion:

Once characteristic which makes the Poisson distribution tractable is that the single parameter  $\lambda$  is the mean and variance of the distribution, i.e.  $E(N) = \lambda = Var(N)$ . Also, the maximum likelihood estimate (*MLE*) of  $\lambda$  is simply the sample mean  $\bar{X} = \sum X_i / N$ , for a sample of size  $N$ .

The Poisson distribution has many additional characteristics which make it favorable for modeling, however one characteristic is closely related to the structure of frequency contagion models, in this paper. Within the Bayesian modeling framework, the parameter  $\lambda$  of the Poisson distribution is considered to be a RV. In particular, if  $\lambda$  is assumed to follow a Gamma distribution of the following form:

$$f(\lambda) = \frac{1}{\Gamma(\alpha)\theta^\alpha} \lambda^{\alpha-1} e^{-\lambda/\theta} \quad \text{for } \alpha, \theta > 0$$

where  $\alpha$  is the shape parameter, and  $\theta$  is the scale parameter, then the marginal distribution, or predictive distribution, of  $N$  can be shown to be a Negative Binomial distribution with parameters  $r = \alpha$  and  $p = 1/(1 + \theta)$ .

Note that this Bayesian, Poisson-Gamma, model is often depicted, symbolically, as:  $N|\lambda \sim \text{Poisson}(\lambda)$  with  $\lambda \sim \text{Gamma}(\alpha, \theta)$ , and is described as *mixing* a Poisson with a gamma. Further, since  $\lambda$  is considered a RV, the distribution of  $N$  depends on the value of the RV  $\lambda$ , hence the notation  $N|\lambda$ , which denotes the conditional value of  $N$ , given the value of  $\lambda$ . Using elementary probability theory, including Bayes' theorem, the marginal distribution, or predictive distribution, of  $N$  is determined by the following integral:

$$Pr(N = n) = \int_0^\infty pr(N = n | \lambda) f(\lambda) d\lambda$$

The mean of the resulting negative binomial distribution is:  $E(N) = r(1 - p)/p = r\theta$ , and the variance is:  $Var(N) = r\theta(1 + \theta)$ . Notice that the variance can be written as:

$$Var(N) = E(N)(1 + \theta)$$

It can be seen that the variance of the negative binomial distribution is larger (by a factor of  $1 + \theta$ ) than the mean. Hence, by starting with a distribution with its mean equal to its variance, and allowing its (only) parameter to be uncertain, the resulting distribution's variance is larger than its mean. This larger variance includes the extra uncertainty created by allowing the parameter to be random.

Lastly note that, based on the formula for the variance, we have:

$$Var(N)/E(N) = 1 + \theta$$

Frequency contagion models incorporate the extra uncertainty caused by changing underwriting, claims handling, and economic environments, in a similar way.

## 9.2 The Negative Binomial Distribution:

The Negative Binomial distribution is a discrete, two-parameter, distribution with parameters often denoted by  $r$  and  $p$ . A common application of the Negative Binomial is to modeling the number of failures before the  $r^{th}$  success in a series of independent observations, each with the same probability of a success,  $p$ . This parameterization of the Negative Binomial distribution is illustrated as follows:

$$\Pr(N = n) = \binom{r+n-1}{n} p^r (1-p)^n \quad \text{for } n \geq 0 \text{ and } 0 < p < 1, r \in \mathbb{Z}^+$$

Under this parameterization, it can be shown that the mean and variance of the Negative Binomial are:  $E(N) = r(1-p)/p$  and  $Var(N) = r(1-p)/p^2$

Since,  $0 < p < 1$ , the variance of the Negative Binomial is greater than the mean. This observation is often a key consideration when choosing between the Poisson or Negative Binomial, for a given data set.

The Negative binomial can also be parameterized by letting  $\beta = 1/p - 1$ . Under this parameterization we have:

$$\Pr(N = n) = \binom{r+n-1}{n} \left(\frac{1}{1+\beta}\right)^r \left(\frac{\beta}{1+\beta}\right)^n \quad \text{for } n \geq 0 \text{ and } \beta > 0, r \in \mathbb{Z}^+$$

The mean and variance can be written, somewhat more simply, under this parameterization as:

$$E(N) = r\beta \quad \text{and} \quad Var(N) = r\beta(1 + \beta)$$

As before, to denote that  $N$  follows a Negative Binomial distribution, with parameters  $r$  and  $\beta$ , we write:  $N \sim \text{NegBin}(r, \beta)$ .

The Contagion modeling framework can also be applied when claim counts are best modeled with a Negative Binomial distribution, rather than a Poisson distribution. However, care must be taken in determining how the contagion RV will be incorporated within the Negative Binomial distribution. As previously discussed, the resulting model should have several attributes, including that the mean of the resulting model be unaffected by the introduction of the contagion RV, and also that the contagion parameter,  $c$ , influences the resulting model in a way which is

consistent with the theory underlying contagion. This is complicated by the fact that the Negative Binomial is a two-parameter distribution, and hence there are more potential ways to introduce the contagion RV.

### 9.3 Variance of Negative Binomial frequency Contagion RV:

Assuming  $N|C \sim \text{NegBin}(C\lambda, \gamma)$ , and  $E[C] = 1$ , and  $\text{Var}[C] = c$ :

$$\begin{aligned}\text{Var}(N) &= E_C[\text{Var}(N|C)] + \text{Var}_C[E(N|C)] = \\ &= E_C[C\lambda + (C\lambda)^2\gamma] + \text{Var}_C(C\lambda) = \\ &= \lambda + \lambda^2\gamma(1 + c) + \lambda^2c = \\ &= \lambda(1 + \lambda(c + c\gamma + \gamma))\end{aligned}$$

### 9.4 Covariance between Negative Binomial frequency Contagion RV's:

$$\begin{aligned}\text{Cov}(N_i, N_j) &= E(N_i N_j) - E(N_i)E(N_j) = \\ &= E(C^2 \lambda_i \lambda_j) - \lambda_i \lambda_j = \\ &= \lambda_i \lambda_j E(C^2) - \lambda_i \lambda_j = \\ &= c\lambda_i \lambda_j + \lambda_i \lambda_j - \lambda_i \lambda_j = \\ &= c\lambda_i \lambda_j,\end{aligned}$$

for  $1 \leq i, j \leq K$ , and  $i \neq j$ .

### 9.5 Correlation between Negative Binomial frequency Contagion RV's:

$$\begin{aligned}\rho_{N_i, N_j} &= \frac{c\lambda_i \lambda_j}{\sqrt{\lambda_i(1 + \lambda_i(c + c\gamma_i + \gamma_i)) \cdot \lambda_j(1 + \lambda_j(c + c\gamma_j + \gamma_j))}} = \\ &= \sqrt{\frac{c\lambda_1}{1 + \lambda_1(c + c\gamma_1 + \gamma_1)}} \sqrt{\frac{c\lambda_2}{1 + \lambda_2(c + c\gamma_2 + \gamma_2)}},\end{aligned}$$

for  $1 \leq i, j \leq K$ , and  $i \neq j$ .

### 9.6 Mean of Binomial frequency Contagion RV:



$$\begin{aligned}
E(N) &= E_p[E_N(N | p)] = E_p[\hat{n}p] = \hat{n} \cdot E_p[p] = \\
&= \hat{n} \cdot \frac{\frac{1}{c}}{\frac{1}{c} + \frac{1}{c} \left( \frac{1-\hat{p}}{\hat{p}} \right)} = \hat{n} \cdot \hat{p}
\end{aligned}$$

## 9.7 Variance of Binomial frequency Contagion RV:

To calculate the variance, first note that:

$$\begin{aligned}
E[p] &= \frac{\alpha}{\alpha + \beta} = \frac{\frac{1}{c}}{\frac{1}{c} + \frac{1-\hat{p}}{\hat{p}} \left( \frac{1}{c} \right)} = \hat{p} \\
Var[p] &= \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = \frac{\frac{1}{c^2} \left( \frac{1-\hat{p}}{\hat{p}} \right)}{\frac{1}{c^2 \hat{p}^2} \left( \frac{1+c\hat{p}}{c\hat{p}} \right)} = \frac{c\hat{p}^2(1-\hat{p})}{1+c\hat{p}}
\end{aligned}$$

And:

$$E[p^2] = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} = \frac{\left( \frac{1}{c} \left( \frac{1}{c} + 1 \right) \right)}{\frac{1}{c\hat{p}} \left( \frac{1}{c\hat{p}} + 1 \right)} = \frac{\hat{p}^2(1+c)}{1+c\hat{p}}$$

So that:

$$\begin{aligned}
E_p[Var(N|p)] &= E_p[\hat{n}p(1-p)] = E_p[\hat{n}p] - E_p[\hat{n}p^2] = \hat{n} \cdot E_p[p] - \hat{n} \cdot E_p[p^2] = \\
&= \hat{n}\hat{p} - \hat{n} \frac{\hat{p}^2(1+c)}{1+c\hat{p}} = \frac{\hat{n}\hat{p}(1+c\hat{p})}{1+c\hat{p}} - \hat{n} \frac{\hat{p}^2(1+c)}{1+c\hat{p}} = \\
&= \frac{(\hat{n}\hat{p} + c\hat{n}\hat{p}^2) - (\hat{n}\hat{p}^2 + c\hat{n}\hat{p}^2)}{1+c\hat{p}}.
\end{aligned}$$

And further:

$$Var_p[E(N|p)] = \hat{n}^2 Var_p[p] = \hat{n}^2 \frac{c\hat{p}^2(1-\hat{p})}{1+c\hat{p}}.$$

Therefore:

$$\begin{aligned}
Var(N) &= E_p[Var(N|p)] + Var_p[E(N|p)] = \\
&= \left( \frac{(\hat{n}\hat{p} + c\hat{n}\hat{p}^2) - (\hat{n}\hat{p}^2 + c\hat{n}\hat{p}^2)}{1 + c\hat{p}} \right) + \hat{n}^2 \frac{c\hat{p}^2(1 - \hat{p})}{1 + c\hat{p}} = \\
&= \frac{(\hat{n}\hat{p} + c\hat{n}\hat{p}^2) - (\hat{n}\hat{p}^2 + c\hat{n}\hat{p}^2) + \hat{n}^2 c\hat{p}^2(1 - \hat{p})}{1 + c\hat{p}} = \\
&= \frac{\hat{n}\hat{p}(1 - \hat{p}) + \hat{n}^2 c\hat{p}^2(1 - \hat{p})}{1 + c\hat{p}} = \hat{n}\hat{p}(1 - \hat{p}) \cdot \frac{1 + \hat{n}\hat{p}c}{1 + \hat{p}c}.
\end{aligned}$$

9.8 Mean of Binomial frequency Contagion RV, for the  $i^{th}$  out of  $K$  lines of business:

$$\begin{aligned}
E(N_i) &= E_p[E_{N_i}(N_i | p)] = E_p \left[ \hat{n}_i \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \right] = \hat{n}_i \left( \frac{\hat{p}_i}{\hat{p}^*} \right) \cdot E_p[p] = \\
&= \hat{n}_i \left( \frac{\hat{p}_i}{\hat{p}^*} \right) \cdot \frac{\frac{1}{c}}{\frac{1}{c} + \frac{1 - \hat{p}^*}{\hat{p}^*} \left( \frac{1}{c} \right)} = \hat{n}_i \left( \frac{\hat{p}_i}{\hat{p}^*} \right) \cdot \hat{p}^* = \hat{n}_i \cdot \hat{p}_i
\end{aligned}$$

9.9 Variance of Binomial frequency Contagion RV, for the  $i^{th}$  out of  $K$  lines of business:

Since:  $Var(N_i) = E[Var(N_i|p)] + Var[E(N_i|p)]$  we need  $E[Var(N_i|p)]$  and  $Var[E(N_i|p)]$ :

$$\begin{aligned}
E_p[Var(N|p)] &= E_p \left[ \hat{n} \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \left( 1 - \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \right) \right] = E_p \left[ \hat{n} \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \right] - E_p \left[ \hat{n} \left( \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \right)^2 \right] = \\
&= \hat{n} \left( \frac{\hat{p}_i}{\hat{p}^*} \right) \cdot E_p[p] - \hat{n} \left( \frac{\hat{p}_i}{\hat{p}^*} \right)^2 \cdot E_p[p^2].
\end{aligned}$$

And:

$$Var_p[E(N_i|p)] = Var_p\left[\hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right)p\right] = \left(\hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right)\right)^2 Var_p[p] = \hat{n}_i^2\left(\frac{\hat{p}_i}{\hat{p}^*}\right)^2 Var_p[p].$$

Again, as for a single Binomial, we need  $E[p]$ ,  $E[p^2]$ , and  $Var[p]$ , in order to compute the above quantities. Based on the Beta prior distribution for  $p$ , we have:

$$E[p] = \frac{\alpha}{\alpha + \beta} = \frac{\frac{1}{c}}{\frac{1}{c} + \frac{1 - \hat{p}^*}{\hat{p}^*}\left(\frac{1}{c}\right)} = \hat{p}^*$$

and:

$$Var[p] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)} = \frac{\frac{1}{c^2}\left(\frac{1 - \hat{p}^*}{\hat{p}^*}\right)}{\frac{1}{c^2\hat{p}^{*2}}\left(\frac{1 + c\hat{p}^*}{c\hat{p}^*}\right)} = \frac{c \cdot (\hat{p}^*)^2(1 - \hat{p}^*)}{1 + c\hat{p}^*}$$

and:

$$E[p^2] = \frac{\alpha(\alpha + 1)}{(\alpha + \beta)(\alpha + \beta + 1)} = \frac{\left(\frac{1}{c}\left(\frac{1}{c} + 1\right)\right)}{\frac{1}{c\hat{p}^*}\left(\frac{1}{c\hat{p}^*} + 1\right)} = \frac{(\hat{p}^*)^2(1 + c)}{1 + c\hat{p}^*}$$

So that we have:

$$\begin{aligned} E_p[Var(N|p)] &= \hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right) \cdot E_p[p] - \hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right)^2 \cdot E_p[p^2] = \\ &= \hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right) \cdot \hat{p}^* - \hat{n}_i\left(\frac{\hat{p}_i}{\hat{p}^*}\right)^2 \cdot \frac{(\hat{p}^*)^2(1 + c)}{1 + c\hat{p}^*} = \\ &= \hat{n}_i\hat{p}_i - \frac{\hat{n}_i(\hat{p}_i)^2(1 + c)}{1 + c\hat{p}^*}. \end{aligned}$$

and:

$$\begin{aligned} Var_p[E(N_i|p)] &= \hat{n}_i^2\left(\frac{\hat{p}_i}{\hat{p}^*}\right)^2 Var_p[p] = \hat{n}_i^2\left(\frac{\hat{p}_i}{\hat{p}^*}\right)^2 \frac{c \cdot (\hat{p}^*)^2(1 - \hat{p}^*)}{1 + c\hat{p}^*} = \\ &= \frac{c \cdot (\hat{n}_i\hat{p}_i)^2(1 - \hat{p}^*)}{1 + c\hat{p}^*}. \end{aligned}$$

So that, all together:

$$\begin{aligned}
Var(N_i) &= E_p[Var(N_i|p)] + Var_p[E(N_i|p)] = \\
&= \left( \hat{n}_i \hat{p}_i - \frac{\hat{n}_i (\hat{p}_i)^2 (1+c)}{1+c\hat{p}^*} \right) + \frac{c \cdot (\hat{n}_i \hat{p}_i)^2 (1-\hat{p}^*)}{1+c\hat{p}^*} = \\
&= \left( \frac{\hat{n}_i \hat{p}_i (1+c\hat{p}^*) - \hat{n}_i (\hat{p}_i)^2 (1+c)}{1+c\hat{p}^*} \right) + \frac{c \cdot (\hat{n}_i)^2 (\hat{p}_i)^2 (1-\hat{p}^*)}{1+c\hat{p}^*} = \\
&= \frac{\hat{n}_i \hat{p}_i - \hat{n}_i (\hat{p}_i)^2 + c\hat{n}_i (\hat{p}_i) \hat{p}^* - c\hat{n}_i (\hat{p}_i)^2 + c \cdot (\hat{n}_i)^2 (\hat{p}_i)^2 (1-\hat{p}^*)}{1+c\hat{p}^*} = \\
&= \frac{\hat{n}_i \hat{p}_i (1-\hat{p}_i) + c\hat{p}^* \left( \hat{n}_i \hat{p}_i - \hat{n}_i \frac{(\hat{p}_i)^2}{\hat{p}^*} + (\hat{n}_i)^2 \frac{(\hat{p}_i)^2}{\hat{p}^*} (1-\hat{p}^*) \right)}{1+c\hat{p}^*} = \\
&= \frac{\hat{n}_i \hat{p}_i (1-\hat{p}_i) + c\hat{p}^* \left[ \hat{n}_i \hat{p}_i \left( 1 - \frac{\hat{p}_i}{\hat{p}^*} \right) + (\hat{n}_i)^2 (\hat{p}_i)^2 \left( \frac{1}{\hat{p}^*} - 1 \right) \right]}{1+c\hat{p}^*}.
\end{aligned}$$

## 9.10 Proof that $Var(N_i)$ is an increasing function of $c$ , in the Binomial case:

To see that  $Var(N_i)$  is an increasing function of  $c$  we consider two cases;  $\hat{p}_i = \hat{p}^*$  and  $\hat{p}_i < \hat{p}^*$ . In case 1, we assume that  $\hat{p}_i \geq \hat{p}_j$  for all  $0 \leq j \leq K$ , then  $\hat{p}_i = \hat{p}^*$ , and  $Var(N_i) \rightarrow (\hat{n}_i)^2 \hat{p}_i (1-\hat{p}_i) \geq \hat{n}_i \hat{p}_i (1-\hat{p}_i)$ .

On the other hand, i.e. case 2, if  $\hat{p}_i \neq \hat{p}^*$ , then  $\hat{p}_i < \hat{p}^*$ , and 1. and 2. above ensure that  $Var(N_i) > \hat{n}_i \hat{p}_i (1-\hat{p}_i)$ . However, this can also be seen by rewriting the limiting formula for  $Var(N_i)$  as:

$$Var(N_i) \rightarrow \hat{n}_i \hat{p}_i \left( 1 - \frac{\hat{p}_i}{\hat{p}^*} \right) + (\hat{n}_i)^2 (\hat{p}_i)^2 \left( \frac{1}{\hat{p}^*} - 1 \right) = \hat{n}_i \hat{p}_i \left[ 1 - \left( \hat{n}_i \hat{p}_i - \frac{\hat{n}_i \hat{p}_i}{\hat{p}^*} + \frac{\hat{p}_i}{\hat{p}^*} \right) \right]$$

And further, since  $\hat{p}^* \leq 1$ , this implies:

$$\left(1 - \frac{1}{\hat{p}^*}\right) > \hat{n}_i \left(1 - \frac{1}{\hat{p}^*}\right) \rightarrow \hat{p}_i > \hat{n}_i \hat{p}_i - \frac{\hat{n}_i \hat{p}_i}{\hat{p}^*} + \frac{\hat{p}_i}{\hat{p}^*}$$

Hence, we see that:

$$Var(N_i) = \hat{n}_i \hat{p}_i \left[ 1 - \left( \hat{n}_i \hat{p}_i - \frac{\hat{n}_i \hat{p}_i}{\hat{p}^*} + \frac{\hat{p}_i}{\hat{p}^*} \right) \right] \geq \hat{n}_i \hat{p}_i (1 - \hat{p}_i)$$

### 9.11 Covariance between the $i^{th}$ and $j^{th}$ Binomial frequency Contagion RVs:

$$\begin{aligned} Cov(N_i, N_j) &= E[N_i N_j] - E[N_i]E[N_j] = E_p \left[ E[N_i N_j | p] \right] - E[N_i]E[N_j] = \\ &= E \left[ \hat{n}_i \left( \frac{\hat{p}_i}{\hat{p}^*} \right) p \cdot \hat{n}_j \left( \frac{\hat{p}_j}{\hat{p}^*} \right) p \right] - \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j = \frac{\hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j}{(\hat{p}^*)^2} \left( \frac{(\hat{p}^*)^2 (1 + c)}{1 + c \hat{p}^*} \right) - \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j = \\ &= \frac{\hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j}{(\hat{p}^*)^2} \left( \frac{(\hat{p}^*)^2 (1 + c)}{1 + c \hat{p}^*} \right) - \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j = \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \left( \frac{1 + c}{1 + c \hat{p}^*} - 1 \right) = \\ &= \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \left( \frac{1 + c}{1 + c \hat{p}^*} - 1 \right) = \hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \left( \frac{c(1 - \hat{p}^*)}{1 + c \hat{p}^*} \right), \end{aligned}$$

for  $1 \leq i, j \leq K$ , and  $i \neq j$ .

### 9.12 Correlation between the $i^{th}$ and $j^{th}$ Binomial frequency Contagion RVs, as $c \rightarrow \infty$ :

For  $1 \leq i, j \leq K$  and  $i \neq j$ , as  $c \rightarrow \infty$ , we have:

$$\rho_{N_i, N_j} \rightarrow \frac{\hat{n}_i \hat{p}_i \hat{n}_j \hat{p}_j \frac{(1 - \hat{p}^*)}{\hat{p}^*}}{\sqrt{\hat{n}_i \hat{p}_i \left[ 1 - \left( \hat{n}_i \hat{p}_i - \frac{\hat{n}_i \hat{p}_i}{\hat{p}^*} + \frac{\hat{p}_i}{\hat{p}^*} \right) \right]} \sqrt{\hat{n}_j \hat{p}_j \left[ 1 - \left( \hat{n}_j \hat{p}_j - \frac{\hat{n}_j \hat{p}_j}{\hat{p}^*} + \frac{\hat{p}_j}{\hat{p}^*} \right) \right]}} =$$

$$\begin{aligned}
&= \frac{(1 - \hat{p}^*) \hat{n}_i \hat{n}_j \hat{p}_i \hat{p}_j}{\sqrt{\hat{n}_i \hat{p}_i [\hat{p}^* - (\hat{n}_i \hat{p}_i \hat{p}^* - \hat{n}_i \hat{p}_i + \hat{p}_i)]} \sqrt{\hat{n}_j \hat{p}_j [\hat{p}^* - (\hat{n}_j \hat{p}_j \hat{p}^* - \hat{n}_j \hat{p}_j + \hat{p}_j)]}} = \\
&= \frac{(1 - \hat{p}^*) \hat{n}_i \hat{n}_j \hat{p}_i \hat{p}_j}{\sqrt{(\hat{n}_i \hat{p}_i)^2 \left[1 - \hat{p}^* + \frac{\hat{p}^*}{\hat{n}_i \hat{p}_i} - \frac{1}{\hat{n}_i}\right]} \sqrt{(\hat{n}_j \hat{p}_j)^2 \left[1 - \hat{p}^* + \frac{\hat{p}^*}{\hat{n}_j \hat{p}_j} - \frac{1}{\hat{n}_j}\right]}} = \\
&= \frac{1}{\sqrt{\left(1 + \frac{(\hat{p}^* - \hat{p}_i)}{\hat{n}_i \hat{p}_i (1 - \hat{p}^*)}\right) \left(1 + \frac{(\hat{p}^* - \hat{p}_j)}{\hat{n}_j \hat{p}_j (1 - \hat{p}^*)}\right)}} .
\end{aligned}$$

9.13 Variance of severity for the  $k$ -th line of business, under the severity contagion decomposition method,  $\beta Z_k$ :

$$\begin{aligned}
Var[\beta Z_k] &= Var_{\beta} [E[\beta Z_k | \beta]] + E_{\beta} [Var[\beta Z_k | \beta]] = \\
&= Var_{\beta} [\beta \mu_k] + E_{\beta} [\beta^2 \sigma_{z_k}^2] = \\
&= \mu_k^2 Var_{\beta} [\beta] + \sigma_{z_k}^2 E_{\beta} [\beta^2] = \\
&= \mu_k^2 b + \sigma_{z_k}^2 (1 + b) = \\
&= \sigma_{z_k}^2 + b(\mu_k^2 + \sigma_{z_k}^2).
\end{aligned}$$

9.14 Covariance between the severities for the  $k^{th}$  and  $j^{th}$  lines of business, under the severity contagion decomposition method,  $\beta Z_k$ , and  $\beta Z_j$ :

$$\begin{aligned}
 Cov(\beta Z_k, \beta Z_j) &= E(\beta^2 Z_k Z_j) - [E(\beta)]^2 E(Z_k) E(Z_j) = \\
 &= E(\beta^2) E(Z_k Z_j) - [E(\beta)]^2 E(Z_k) E(Z_j) = \\
 &= E(\beta^2) E(Z_k) E(Z_j) - [E(\beta)]^2 E(Z_k) E(Z_j) = \\
 &= E(Z_k) E(Z_j) [E(\beta^2) - [E(\beta)]^2] = \\
 &= E(Z_k) E(Z_j) Var(\beta) = \\
 &= \mu_{z_k} \mu_{z_j} b = \\
 &= \mu_k \mu_j b.
 \end{aligned}$$

9.15 Correlation between the severities for the  $k^{th}$  and  $j^{th}$  lines of business, under the severity contagion decomposition method,  $\beta Z_k$ , and  $\beta Z_j$ :

Thus, the correlation between  $\beta Z_k$ , and  $\beta Z_j$ , is:

$$\begin{aligned}
 \rho_{\beta Z_k, \beta Z_j} &= \frac{Cov(\beta Z_k, \beta Z_j)}{\sqrt{Var(\beta Z_k) Var(\beta Z_j)}} = \\
 &= \frac{\mu_k \mu_j b}{\sqrt{[\sigma_{z_k}^2 + b(\mu_k^2 + \sigma_{z_k}^2)] [\sigma_{z_j}^2 + b(\mu_j^2 + \sigma_{z_j}^2)]}} = \\
 &= \frac{\mu_k \mu_j b}{\sqrt{\mu_k^2 \left[ \frac{\sigma_{z_k}^2}{\mu_k^2} + b \left( 1 + \frac{\sigma_{z_k}^2}{\mu_k^2} \right) \right] \mu_j^2 \left[ \frac{\sigma_{z_j}^2}{\mu_j^2} + b \left( 1 + \frac{\sigma_{z_j}^2}{\mu_j^2} \right) \right]}} = \\
 &= \frac{b}{\sqrt{[b + (1 + b) CV_{z_k}] [b + (1 + b) CV_{z_j}]}}
 \end{aligned}$$

### 9.16 Expected value of the aggregate losses for a single line, under both frequency and severity Contagion:

$$\begin{aligned}
 E[S^*] &= E_{N^*} \left( E \left[ \sum_{i=1}^{N^*} \beta Z_i \middle| N^* \right] \right) = E_{N^*} [N^* \cdot E[\beta Z_i]] = E_{N^*} [N^*] E[\beta Z_i] = E_{N^*} [N^*] (\beta \cdot E[Z_i]) = \\
 &= \lambda \cdot \mu
 \end{aligned}$$

### 9.17 Variance of the aggregate losses for a single line, under both frequency and severity Contagion:

To calculate the variance of the aggregate loss, under both frequency and severity contagion, first note that:

$$E[S^*|\beta] = E_{N^*} \left( E \left[ \sum_{i=1}^{N^*} \beta Z_i \middle| N^*, \beta \right] \right) = E_{N^*} (N^* \beta E[Z_i]) = \beta \cdot E[Z_i] E_{N^*} (N^*).$$

and:

$$\begin{aligned}
 Var[S^*|\beta] &= Var_{N^*} \left( E \left[ \sum_{i=1}^{N^*} \beta Z_i \middle| N^*, \beta \right] \right) + E_{N^*} \left( Var \left[ \sum_{i=1}^{N^*} \beta Z_i \middle| N^*, \beta \right] \right) = \\
 &= Var_{N^*} (N^* \beta \cdot E[Z_i]) + E_{N^*} (N^* \beta^2 Var[Z_i]) = \\
 &= \beta^2 (E[Z_i])^2 \cdot Var_{N^*} (N^*) + \beta^2 Var[Z_i] E_{N^*} (N^*) = \\
 &= \beta^2 \{ (E[Z_i])^2 \cdot Var_{N^*} (N^*) + Var[Z_i] E_{N^*} (N^*) \}.
 \end{aligned}$$

And, hence:

$$\begin{aligned}
 Var[S^*] &= E_{\beta} [Var(S^*|\beta)] + Var_{\beta} [E(S^*|\beta)] = \\
 &= E_{\beta} [\beta^2 \{ (E[Z_i])^2 \cdot Var_{N^*} (N^*) + Var[Z_i] E_{N^*} (N^*) \}] + Var_{\beta} [\beta \cdot E[Z_i] E_{N^*} (N^*)] =
 \end{aligned}$$



$$\begin{aligned}
&= E_\beta [\beta^2] \cdot \{(E[Z_i])^2 \cdot \text{Var}_{N^*}(N^*) + \text{Var}[Z_i]E_{N^*}(N^*)\} + (E[Z_i])^2 (E(N^*))^2 \text{Var}_\beta [\beta \cdot] = \\
&= (b + 1) \cdot \{(E[Z_i])^2 \cdot \text{Var}_{N^*}(N^*) + \text{Var}[Z_i]E_{N^*}(N^*)\} + (E[Z_i])^2 (E(N^*))^2 \cdot b = \\
&= \text{Var}[Z_i]E(N^*) + (E[Z_i])^2 \text{Var}(N^*) + b \cdot \{\text{Var}[Z_i]E(N^*) + (E[Z_i])^2 E(N^{*2})\}.
\end{aligned}$$

### 9.18 Covariance between the aggregate losses of two lines of business, under both frequency and severity Contagion:

To compute  $\text{Cov}(S_k^*, S_j^*)$ , for  $k \neq j$ , note that:

$$E[S_k^* S_j^*] = E_{N_k^*}(N_k^*) E[Z_{ki}] \cdot E_{N_j^*}(N_j^*) E[Z_{ji}] \cdot E[\beta^2]$$

Since:

$$\begin{aligned}
E_{S_k^* \times S_j^*}[S_k^* S_j^* | \beta] &= E_{N_k^* \times N_j^*} \left( E_{S_k^* \times S_j^*} \left[ \left( \sum_{i=1}^{N_k^*} \beta Z_{ki} \right) \left( \sum_{i=1}^{N_j^*} \beta Z_{ji} \right) \middle| \beta, N_j^*, N_k^* \right] \right) = \\
&= E_{N_k^* \times N_j^*} (N_k^* \beta \cdot N_j^* \beta \cdot E_{Z_k \times Z_j}[Z_{ki} Z_{ji}]) = \\
&= E_{N_k^* \times N_j^*} (N_k^* \beta \cdot N_j^* \beta \cdot E[Z_{ki}] E[Z_{ji}]) = \\
&= \beta^2 E[Z_{ki}] E[Z_{ji}] \cdot E_{N_k^* \times N_j^*}(N_k^* N_j^*).
\end{aligned}$$

And hence:

$$E[S_k^* S_j^*] = E_\beta [\beta^2 E[Z_{ki}] E[Z_{ji}] \cdot E_{N_k^* \times N_j^*}(N_k^* N_j^*)] = E[Z_{ki}] E[Z_{ji}] \cdot E(N_k^* N_j^*) \cdot E[\beta^2].$$

Therefore, using  $E[S^*] = E_{N^*}[N^*] E[Z_i]$ , from above:

$$\begin{aligned}
\text{Cov}(S_k^*, S_j^*) &= E[S_k^* S_j^*] - E[S_k^*] E[S_j^*] = \\
&= E[Z_{ki}] E[Z_{ji}] \cdot E(N_k^* N_j^*) \cdot E[\beta^2] - E(N_k^*) E[Z_{ki}] E[N_j^*] E[Z_{ji}] =
\end{aligned}$$

$$\begin{aligned}
&= E[Z_{ki}]E[Z_{ji}] \cdot \{E(N_k^* N_j^*) \cdot E[\beta^2] - E(N_k^*)E[N_j^*]\} = \\
&= E[Z_{ki}]E[Z_{ji}] \cdot \{E(N_k^* N_j^*) \cdot (b + 1) - E(N_k^*)E[N_j^*]\} = \\
&= E[Z_{ki}]E[Z_{ji}] \cdot \{b \cdot E(N_k^* N_j^*) + Cov(N_k^*, N_j^*)\}.
\end{aligned}$$

### 9.19 Correlation between the aggregate losses of two lines of business, under both frequency and severity Contagion:

Using that we can write:  $Var[S^*] = \Sigma + b \cdot [\Sigma + (\mu\lambda)^2]$

Where:

- $\Sigma = Var[Z_i]E(N^*) + (E[Z_i])^2 Var(N^*)$
- $\mu = E[Z_i]$
- $\lambda = E(N) = E(N^*)$

$$\begin{aligned}
\rho_{S_k^*, S_j^*} &= \frac{Cov(S_k^*, S_j^*)}{\sqrt{Var[S_k^*]Var[S_j^*]}} = \frac{\mu_k \mu_j \cdot \{b \cdot E(N_k^* N_j^*) + Cov(N_k^*, N_j^*)\}}{\sqrt{(\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]) (\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2])}} = \\
&= \frac{\lambda_k \mu_k \lambda_j \mu_j \cdot (cb + b + c)}{\sqrt{(\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]) (\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2])}}.
\end{aligned}$$

## 9.20 Correlation between the aggregate losses – special case Poisson Frequencies:

In the specific case of Poisson Frequency contagion, the  $\Sigma$  within Equation (5.10) can be written:

$$\begin{aligned}\Sigma &= \text{Var}[Z_i]E(N^*) + (E[Z_i])^2\text{Var}(N^*) = \lambda \text{Var}[Z_i] + (E[Z_i])^2[\lambda(1 + c\lambda)] = \\ &= \lambda \text{Var}[Z_i] + \lambda(E[Z_i])^2 + c\lambda^2(E[Z_i])^2.\end{aligned}$$

However, we can recognize  $\lambda \text{Var}[Z_i] + \lambda(E[Z_i])^2$  as the variance Aggregate losses under the Traditional Collective Risk model, under Poisson Frequency, and using  $Z_i$  as the severity  $RV$ , ie:

$$\text{Var}(S) = \lambda \text{Var}[Z_i] + \lambda(E[Z_i])^2$$

Where:

$$S = \sum_{i=1}^N Z_i$$

And:

$$\begin{cases} N \sim \text{Poisson}(\lambda) \\ Z_i \sim \text{Dist}_Z(\mu, \sigma_Z^2) \end{cases}$$

Therefore:

$$\Sigma = \text{Var}(S) + c\lambda^2(E[Z_i])^2 = \text{Var}(S) + c\lambda^2\mu^2$$

So:

$$\begin{aligned}\text{Var}[S^*] &= \text{Var}(S) + c\lambda^2\mu^2 + b \cdot [\text{Var}(S) + c\lambda^2\mu^2 + (\mu\lambda)^2] = \\ &= (1 + b)\text{Var}(S) + c\lambda^2\mu^2 + b \cdot (\mu\lambda)^2[c + 1] = \\ &= (1 + b)\text{Var}(S) + (\mu\lambda)^2[bc + b + c].\end{aligned}$$

Therefore, we have that:

$$\rho_{S_k^*, S_j^*} = \frac{\text{Cov}(S_k^*, S_j^*)}{\sqrt{\text{Var}[S_k^*]\text{Var}[S_j^*]}} = \frac{\lambda_k \mu_k \lambda_j \mu_j \cdot (cb + b + c)}{\sqrt{(\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]) (\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2])}} =$$

$$\begin{aligned}
&= \frac{(cb + b + c)}{\sqrt{\frac{1}{(\lambda_k \mu_k)^2} (\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]) \frac{1}{(\lambda_j \mu_j)^2} (\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2])}} = \\
&= \frac{(cb + b + c)}{\sqrt{\left( \frac{\Sigma_k + b \cdot [\Sigma_k + (\mu_k \lambda_k)^2]}{(\lambda_k \mu_k)^2} \right) \left( \frac{\Sigma_j + b \cdot [\Sigma_j + (\mu_j \lambda_j)^2]}{(\lambda_j \mu_j)^2} \right)}} = \\
&= \frac{(cb + b + c)}{\sqrt{\left( \frac{(1+b)Var(S_k) + (\mu_k \lambda_k)^2 [bc + b + c]}{(E[S_k^*])^2} \right) \left( \frac{(1+b)Var(S_j) + (\mu_j \lambda_j)^2 [bc + b + c]}{(E[S_j^*])^2} \right)}} = \\
&= \frac{(cb + b + c)}{\sqrt{\left[ (1+b)CV_{S_k} + [bc + b + c] \right] \left[ (1+b)CV_{S_j} + [bc + b + c] \right]}}
\end{aligned}$$

Or:

$$\rho_{S_k^*, S_j^*} = \frac{(cb + b + c)}{\sqrt{\left[ (1+b)CV_{S_k}^2 + [bc + b + c] \right] \left[ (1+b)CV_{S_j}^2 + [bc + b + c] \right]}}$$

### 9.21 The Pareto (Type II) distribution:

We use the following parameterization of the Pareto distribution:

$$F(x) = \Pr(X \leq x) = 1 - \left[1 + \frac{x}{\theta}\right]^{-\alpha}$$

Where  $x \geq 0$ , and  $\alpha > 0$  is the *shape* parameter, and  $\theta > 0$  is the *scale* parameter.

## 10 References

Klugman, P. W. (2012). *Loss Models - From Data to Decisions* (Forth Edition ed.). Wiley Publishing.

Meyers, G. G. (1997). *The Common Shock Model for Correlated Insurance Losses*. *Variance*, 1(1).

White, S. (2012). *Advanced Correlations*. MetaRisk Conference. Chicago.