1 Introduction

Often the Loss Distribution Approach (LDA) to operational risk assumes that the frequency and severity of operational loss events are independent. The Basel Committee observes that most banks model frequency and severity separately, with the frequency distribution usually modeled by the Poisson or Negative Binomial distributions [2]. This assumption is typically made for computational or estimation purposes. However, it is not difficult to envisage scenarios where this assumption breaks down in practice. Following a particularly severe loss event, the manpower required to deal with the event may cause a failure in internal control leading to an increased likelihood of an additional event. The occurrence of an event may be the result of declining internal controls, which may indicate an increased likelihood of additional events. Because of the possibilities of correlation between the frequency and severity of operational loss events, there is no shortage of papers which test the assumption. Several authors (Brechmann et. al. [4], Mittnick et. al. [14]) examine the correlation of frequency across business lines. Cope and Antonini [8] examine many correlations including correlation between loss event severities. As a result of the empirical studies around correlation, there has been increased attempts to create models which incorporate correlations. The Federal Reserve recently released guidance around operational modeling and included a section on diversification modeling [3]. While making no recommendations, the Federal Reserve mentions copula modeling for dependence. Frachot et. al. [11] attempt to model both correlation between severity and frequency as well as severity auto-correlation, though the model is a single period model which lacks analytic tractability.

I propose a model that treats the operational risk loss variable as a compound Poisson process. By using Carr and Wu's [7] concept of a "leverage neutral" measure I obtain a semi-closed form solution for the characteristic function of the loss distribution while retaining the following key features: the frequency distribution is auto-correlated and the frequency distribution and severity distributions are correlated. This is achieved by letting the process that describes the frequency distribution jump simultaneously with a severity event. A severity event hence increases the likelihood of another severity event occurring in any fixed time period following original event. The standard LDA model is recovered as a special case of this model.

2 Model Description

Consider a business with operational loss events. The random variable that counts the number of loss events in a time period $[0, t]$ is N_t and follows a Poisson distribution with parameter λt. The total loss from these loss events is $X_t = \sum_{j=1}^{N_t} L_j$ where each L_j is independently and identically distributed positive random variables with finite expectation. These L_j represent the dollar loss from loss event j. Since each L_j is stochastic, X_t is a compound Poisson jump process. Intuitively, loss events arrive as a "stream" instead of at a point in time, with no two events occurring simultaneously.

A Poisson jump process's characteristic function is derived as follows:

$$\begin{aligned}
\mathbb{E}[e^{uiX_t}] &= \mathbb{E}\left[\mathbb{E}[e^{uiX_t}|\mathcal{F}_t]\right] = \mathbb{E}\left[\mathbb{E}\left[\prod_{j=1}^{N_t} e^{uiL_j}|\mathcal{F}_t\right]\right] \\
&= \mathbb{E}\left[\prod_{j=1}^{N_t} \mathbb{E}[e^{uiL}]\right] = \mathbb{E}\left[e^{\ln\left(\mathbb{E}[e^{uiL}]\right)N_t}\right] \\
&= e^{\lambda t\left(\mathbb{E}\left[e^{iuL}\right]-1\right)} = e^{t\int_{\mathbb{R}}(e^{iuL}-1)\lambda d\mu}
\end{aligned}$$

Where the second line is justified by the moment generating function of a Poisson random variable, $\mathcal{F}_t$ is the filtration generated by N_t, and $d\mu$ is the density function of L. From the characteristic function, it is clear that the jump diffusion is defined by the product $\lambda d\mu$ where λ controls the frequency of jumps and $d\mu$ controls the jump size. Making this dependence explicit, the characteristic function of a compound Poisson process is denoted

$$\phi_P(u;\, \lambda d\mu, t) = e^{t\int_{\mathbb{R}}(e^{iuL}-1)\lambda d\mu}$$

Comparing the characteristic function to the Levy-Khintchine representation, it is clear that X_t is a Levy process.

From the characteristic function if t is itself random (for notational convenience, τ) and independent of L_j then the extended characteristic function would be

$$\phi_X(u) = \mathbb{E}[e^{uiX_\tau}] = \theta_\tau\left(\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right)\right) \tag{1}$$

Where θ_τ is the moment generating function of τ.

A general specification of τ that admits a (semi) analytic moment generating function is the affine jump-diffusion of Duffie et. al [9]. In this paper, τ will be driven by a particular form of this general jump-diffusion. However, the independence assumption between τ and L is dropped. This leads to a model that has increased tail risk to account for dependencies between the frequency and severity distributions. The goal is to find a semi analytic solution to the characteristic function $\phi_X(u)$.

2.1 Jump specification

For this paper, the jump-diffusion that governs τ is specified as

$$\tau = \int_0^t v(s)ds \tag{2}$$

$$v_t = v_0 + \int_0^t a(1 - \delta\lambda\mathbb{E}[L] - v_s)ds + \sigma \int_0^t \sqrt{v_s}dW_s + \delta \sum_{j=1}^{N_\tau} L_j \tag{3}$$

From here on, δ will be defined as $\delta = \frac{\rho}{\lambda\mathbb{E}[L]}$. This specification allows ρ to have the interpretation as a correlation parameter: it is constrained between [0, 1) and controls the level of correlation between L and λ. Letting $\bar{b} = 1 - \rho$,

$$v_t = v_0 + \int_0^t a(\bar{b} - v_s)ds + \sigma \int_0^t \sqrt{v_s}dW_s + \delta \sum_{j=1}^{N_\tau} L_j \tag{4}$$

Note that N_τ is the time changed counting process N_t and that the long run expected value of v_t is one. Hence the stochastic time is an unbiased estimator of real time. Note also that if L_j is strictly positive then v is positive. Finally, the "standard" LDA model (defined as $\phi_X(u) = e^{t\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right)}$) is recovered when $a = \rho = \sigma = 0$.

This jump-diffusion process for v provides an auto-correlated frequency of jumps ("clumps" of jumps) and correlation between the jump size and the frequency of jumps through the jumps' effect on the level of the process v. Hence this specification fully incorporates correlation between jump size and frequency of jumps and frequency of jumps with itself. Since there is clear correlation between the jump size and jump frequency, equation 1 can no longer be directly applied. However, Carr and Wu [7] showed that by using the "leverage neutral" measure that the characteristic function of X_τ retains analytic tractability under this specification. Their result is used in the next section.

2.2 Semi-analytic solution

The characteristic function of X_τ is

$$\begin{aligned} \mathbb{E}[e^{uiX_\tau}] &= \mathbb{E}\left[e^{uiX_\tau + \tau\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right) - \tau\left(\mathbb{E}\left[e^{iuL}\right]-1\right)}\right] \\ &= \hat{\mathbb{E}}\left[e^{\tau\lambda\mathbb{E}\left[e^{iuL}-1\right]}\right] \end{aligned}$$

Where $\hat{\mathbb{P}}$ is the measure induced by $\eta_\tau = e^{uiX_\tau - \tau\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right)}$.

Theorem 1. Let Y_t be a compound Poisson process characterized by $\lambda d\mu$. Define the martingale

$$\eta_t = e^{uiY_t - t\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right)} \tag{5}$$

Then under the probability measure $d\hat{\mathbb{P}} = \eta_t d\mathbb{P}$, Y_t is a compound Poisson process characterized by $e^{uiL}\lambda d\mu$.

Sketch of Proof.

$$\hat{\mathbb{E}}\left[e^{ziY_t}\right] = \mathbb{E}\left[e^{Y_t i(u+z) - t\lambda\left(\mathbb{E}\left[e^{iuL}\right]-1\right)}\right]$$

$$= e^{t\lambda \mathbb{E}\left[e^{i(z+u)L}\right] - t\lambda \mathbb{E}\left[e^{iuL}\right]}$$
$$= e^{t\int_R \left(e^{izL}-1\right)\lambda e^{iuL}d\mu} = \phi_P(z;\, e^{uiL}\lambda d\mu, t)$$

$\square$

Carr and Wu [7] prove that this theorem remains applicable with τ substituted for t. This theorem implies that under this new measure, the dynamics of v are as follows:

$$v_t = v_0 + \int_0^t a(\bar{b} - v_s)ds + \sigma \int_0^t \sqrt{v_s}dW_s + \delta \sum_{j=1}^{\hat{N}_\tau} \hat{L}_j \tag{6}$$

By Duffie et. al. [9], for the affine process v_t, the following statement holds:

$$\mathbb{E}\left[e^{-r\int_0^T v_s ds}\right] = e^{\gamma(0,T)+\zeta(0,T)v_0}$$

Where γ, ζ satisfy

$$\begin{cases} \frac{\partial \zeta}{\partial t} = a\zeta + r - \frac{1}{2}\sigma^2\zeta^2 - \int_{\mathbb{R}} \left(e^{\delta\zeta L} - 1\right)\lambda e^{uiL}d\mu & \zeta(T,T) = 0 \\ \frac{\partial \gamma}{\partial t} = -a\bar{b}\zeta & \gamma(T,T) = 0 \end{cases}$$

In this case, $r = \lambda(1 - \mathbb{E}[e^{uiL}])$. The system of ODEs thus reads

$$\begin{cases} \frac{\partial \zeta}{\partial t} = a\zeta + \lambda - \lambda\phi_L(-(u+\delta\zeta)i) - \frac{1}{2}\sigma^2\zeta^2 & \zeta(T,T) = 0 \\ \frac{\partial \gamma}{\partial t} = -a\bar{b}\zeta & \gamma(T,T) = 0 \end{cases} \tag{7}$$

Where $\phi_L(u)$ is the characteristic function of L.

3 Severity Distribution

There is a substantial body of literature (Rachev et. al. [16], Gourier et. al. [12], Kerbl [13]) providing empirical evidence for "fat-tailed" (infinite variance) severity distributions. Additionally, there is evidence that for these distributions Monte Carlo simulations can have very slow convergence properties (Brunner et. al. [5]). Hence I specify the severity distribution as a stable distribution to better fit the data and to use an analytical approach where Monte Carlo performance is poor. The stable distribution is defined by its characteristic function[1]:

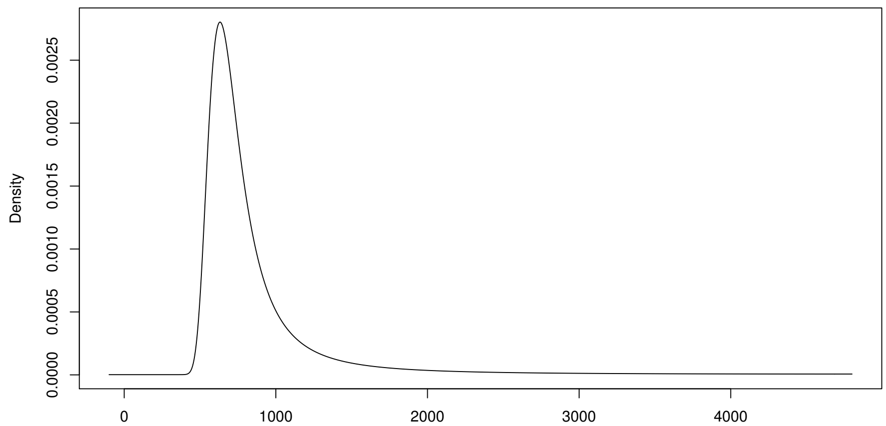
$$\phi_\alpha(u) = e^{iu\mu - (cu)^\alpha\left(1 - \beta i \tan\left(\frac{\pi\alpha}{2}\right)\right)}$$

In general, the stable distribution has support on the reals. When $\beta = 1$ the distribution decays quickly to the left while the right tail remains "fat". The probability of a negative outcome can be made infinitesimal by an appropriate choice of μ and c. To retain finite expectation (a requirement for $\hat{b}$ to be positive)

[1] This characteristic function is valid for $u > 0$ and $\alpha > 1$

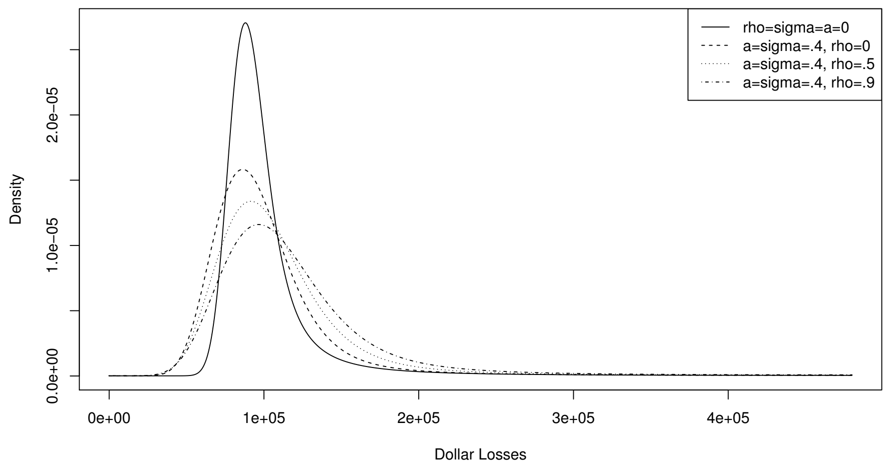
α is constrained to within (1, 2]. The shift parameter μ and the scale parameter c are free parameters which can be calibrated to the historical data.

Figure 1: $\alpha = 1.1$, $\beta = 1$, $c = 100$, $\mu = 1300$



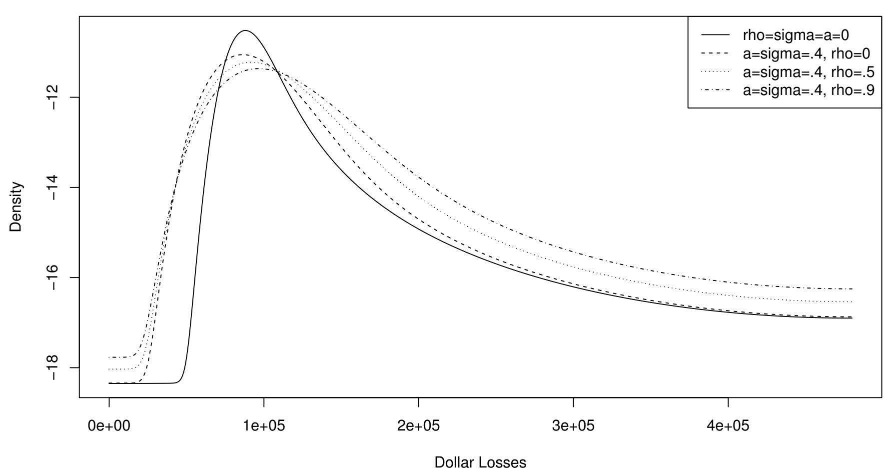
This distribution was leveraged by Carr and Wu [6] to model asset returns: with $\beta = \pm 1$ (maximum skewness) the infinite variance results from thick tails in only one direction. Hence option pricing (which requires finite variance of the asset) and severity modeling (which requires positive severities) can be safely modeled with the stable distribution with maximum skewness.

Figure 2: $\alpha = 1.1$, $\beta = 1$, $c = 100$, $\mu = 1300$, $\lambda = 100$, $t = 1$, ODE steps= 128, X steps= 1024, U steps= 256



To make clear the fatness of the tail, the following is a plot of the natural logarithm of the densities.

Figure 3: $\alpha = 1.1$, $\beta = 1$, $c = 100$, $\mu = 1300$, $\lambda = 100$, $t = 1$, ODE steps= 128, X steps= 1024, U steps= 256



The 99.9% value at risk is dramatically increasing in ρ:

ρ	0	.5	.9
VaR	$2,429,140$	$2,953,580$	$3,539,870$

4 Numerical Implementation

The equations 7 can be numerically solved using the Runge-Kutta method. This method features fourth order convergence (Press et. al. [15]) resulting in highly accurate results for even fairly large step sizes. The accuracy of the results will be explored in a later section. The Runge-Kutta, like all numerical solutions to ODEs, discretizes the domain of the function. The Runge-Kutta uses finite-difference approximations to the derivative of the function in order to compute the solution. The Runge-Kutta improves over naive solutions like Euler's method by taking the "average" of several function solutions in the area of the discrete step. The scheme is given explicitly as follows:

$$\begin{bmatrix} \zeta_{j+1} \\ \gamma_{j+1} \end{bmatrix} = \begin{bmatrix} \zeta_j \\ \gamma_j \end{bmatrix} + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4$$

Where the k_i are defined as follows:

$$k_1 = h\begin{bmatrix} f(jh, \gamma_j, \zeta_j) \\ g(jh, \gamma_j, \zeta_j) \end{bmatrix}$$
$$k_2 = h\begin{bmatrix} f\left(\left(j+\frac{1}{2}\right)h, \gamma_j + \frac{1}{2}k_1, \zeta_j + \frac{1}{2}k_1\right) \\ g\left(\left(j+\frac{1}{2}\right)h, \gamma_j + \frac{1}{2}k_1, \zeta_j + \frac{1}{2}k_1\right) \end{bmatrix}$$
$$k_3 = h\begin{bmatrix} f\left(\left(j+\frac{1}{2}\right)h, \gamma_j + \frac{1}{2}k_2, \zeta_j + \frac{1}{2}k_2\right) \\ g\left(\left(j+\frac{1}{2}\right)h, \gamma_j + \frac{1}{2}k_2, \zeta_j + \frac{1}{2}k_2\right) \end{bmatrix}$$
$$k_4 = h\begin{bmatrix} f\left(\left(j+1\right)h, \gamma_j + k_3, \zeta_j + k_3\right) \\ g\left(\left(j+1\right)h, \gamma_j + k_3, \zeta_j + k_3\right) \end{bmatrix}$$

Here $h = \frac{t}{n}$ where n is the number of steps in the algorithm, j is the j'th step, and f and g are the following functions:

$$f(t, \gamma, \zeta) = \lambda\phi_L(-(u+\delta\zeta)i) + \frac{1}{2}\sigma^2\zeta^2 - a\zeta - \lambda$$
$$g(t, \gamma, \zeta) = a\bar{b}\zeta$$

4.1 Analysis of Complexity

The complexity for computing the characteristic function for a given u $O(n)$ where n is the number of discrete steps in the Runge-Kutta method. To numerically invert the characteristic function, u must be discretized as well. Hence to compute an array of approximate ϕ_X requires $O(nm)$ where m is the number of discrete steps in the complex domain.

Inverting the characteristic function to recover the density using the method proposed by Fang and Oosterlee [10] requires $O(mq)$ operations where q is the number of steps in real domain. The total complexity is thus $O(m(n+q))$.

The speed using c++ and a single core of a fifth generation Intel core i5 is the following for each n:

	$n = 128$	$n = 32$	$n = 2$
Milliseconds	309	90	15

Even with $n = 128$ the algorithm is responsive; while with $n = 2$ the result appears instant to the human perception. The algorithm proposed by Fang and Oosterlee is embarrassingly parallel and can be trivially made to run on multiple cores when available.

4.2 Analysis of Accuracy

The accuracy of Fang and Oosterlee's algorithm is well documented in their paper [10]. The accuracy of the Runge-Kutta algorithm is thus investigated. The following plots show the accuracy of the density for various n.

Absolute difference between densities with $\alpha = 1.1$, $\beta = 1$, $c = 100$, $\mu = 1300$, $\lambda = 100$, $t = 1$, X steps$= 1024$, U steps$= 256$, $a = .4$, $\sigma = .4$, $\rho = .9$

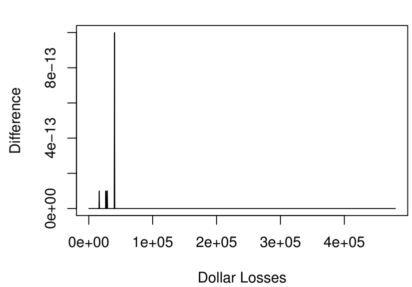


Figure 4: $n = 128$ vs $n = 32$

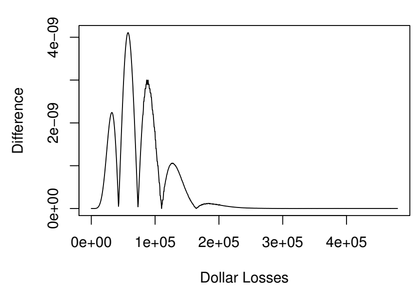


Figure 5: $n = 128$ vs $n = 2$

The difference between the accuracy using $n = 128$ and $n = 32$ is exceedingly small and does not justify the extra computational time. What is perhaps most surprising is that even with $n = 2$ the accuracy is still acceptable especially in the tails.

5 Conclusion

The model presented in this paper provides a method new to the operational risk literature for correlating between severity and frequency in an LDA framework. This model substantially extends the traditional LDA modeling approach by allowing the model to be time-dependent, with stochastic and mean-reverting frequency. The frequency jumps simultaneously with the severity events; inducing correlation between frequency and severity.

While the model substantially generalizes the standard LDA model, there is still room for additional extensions. For example, the Basel committee requires 56 separate operational loss "bins": seven risk types across eight business lines [1]. These risk and business types can be modeled by a multidimensional frequency distribution. In this way the severities across risk and business type will be independent, but a severe event in one business line or risk will directly impact the frequency of severe events in every other risk or business line.

References

[1] BCBS 128: International convergence of capital measurement and capital standards. Technical report, Basel Committee on Banking Supervision, 2006.

[2] BCBS 196: Operational risk - supervisory guidelines for the advanced measurement approaches. Technical report, Basel Committee on Banking Supervision, 2011.

[3] BCC 14-1: Supervisory guidance for data, modeling, and model risk management under the operational risk advanced measurement approaches. Technical report, Federal Reserve Board, 2014.

[4] Eike C. Brechmann, Claudia Cxado, and Sandra Paterlini. Modeling dependence of operational loss frequenecies. *Journal of Operational Risk*, 8, 2013.

[5] Michael Brunner, Fabio Piacenza, Fabio Monti, and Davide Bazzarello. Fat tails, expected shortfall and the monte carlo method: a note. *Journal of Operational Risk*, 4, 2009.

[6] Peter Carr and Liuren Wu. The finite moment log stable process and option pricing. *Journal of Finance*, 58, 2003.

[7] Peter Carr and Liuren Wu. Time-changed levy processes and option pricing. *Journal of Financial Economics*, 71, 2004.

[8] Eric Cope and Gianluca Antonini. Observed correlations and dependencies among operational losses in the orx consortium database. *Journal of Operational Risk*, 3, 2008.

[9] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68, 2000.

[10] F. Fang and C.W. Oosterlee. A novel pricing method for european options based on fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, 31, 2008.

[11] A. Frachot, T. Roncalli, and E. Salomon. The correlation problem in operational risk. *Credit Agricole*, 2004.

[12] Elise Gourier, Walter Farkas, and Donato Abbate. Operational risk quantification using extreme value theory and copulas: from theory to practice. *Journal of Operational Risk*, 4, 2009.

[13] Stefan Kerbl. Evidence, estimates and extreme values from austria. *Journal of Operational Risk*, 9, 2014.

[14] Stefan Mittnik, Sandra Paterlini, and Tina Yener. Modeling dependence of operational loss frequenecies. *Journal of Operational Risk*, 8, 2013.

[15] William Press, Brian Flannery, Saul Teukolsky, and William Vetterling. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge University Press, Cambridge, England, 2 edition, 1992.

[16] Svetlozar Rachev, Anna Chernobai, and Christian Menn. Empirical examination of operational loss distributions. In Martin Morlock, Christoph Schwindt, Norbert Trautmann, and JÃijrgen Zimmermann, editors, *Perspectives on Operations Research*, pages 379–401. Deutscher UniversitÃd'ts-Verlag/GWV Fachverlage GmbH, Wiesbaden, 2006.