APPROXIMATING CORRELATED DEFAULTS FOR CREDIT DEFAULT OPTIONS AND SWAPS

DALE W.R. ROSENTHAL*

Abstract. Modeling defaults is critical to pricing debt portfolio derivatives such as credit default options, collateralized debt obligation tranches and credit default swaps written on those tranches. However, correlated defaults have proven difficult to model. Improper or non-existent modeling of default correlations has caused multi-billion-dollar losses at numerous financial firms. I propose statistical approximations to model correlated defaults. The approximations are consistent and follow from a structural risk factor approach. This approach can price credit default options, such as those currently trading on exchanges, and swaps with an assumption of loss given default. It also yields metrics characterizing portfolio default risk and improving upon a currently-used metric of loan portfolio diversification.
Keywords: correlated defaults, credit default option, CDO, CDS, equity tranche, diversification, diversity score, gamma Edgeworth expansions.
(*JEL*: G13, G12, C16, G33)

* Department of Finance, University of Illinois at Chicago. 601 S. Morgan St. (MC 168), Chicago, IL, USA 60607
I thank my advisor, Per Mykland, and Michael Wichura for helpful comments. Financial support from the National Science Foundation under grants DMS 06-04758 and SES 06-31605 is gratefully acknowledged.

1. Introduction

Many financial phenomena involve waiting for one or more random events. One such phenomenon is default on a loan. To reallocate default risk, many loans might be pooled and the pool tranched — as done with collateralized debt obligations (CDOs). The default risk of a bond or tranche might be hedged with a credit default swap (CDS).

Alternately, one could hedge the default risk on such portfolios by purchasing a credit default option (CDOpt). CDOpts are digital options which pay $\$N$ when a default event occurs and \$0 otherwise. CDOpts are currently traded by the Chicago Board Options Exchange; and, a portfolio of CDOpts at various maturities traded versus a CDS would imply a bet on the term structure of loss given default.

I use an approach based on risk factors to approximate the distribution of correlated default times. This approximation can then be used to price credit default options. Further, if we could model the loss given default as a function of default time, this distribution would allow us to price credit default swaps and collateralized debt obligations.

I assume defaults occur at the expiry of "timers." The idiosyncratic parts of default times are exponentially-distributed. These idiosyncratic timers may be accelerated by the expiry of systematic exponential timers. The interaction of idiosyncratic and systematic timers yields correlated default timers for individual loans.

For example, homeowners default at the expiry of their individual timers; however, their individual timers may be accelerated if the "US recession" timer expires (and a recession ensues).

Mathematically, this structural approach can be recast so that the average default distribution (and preceding metrics) involve sums of exponential random variables. Equally-weighted sums of iid exponential random variables are gamma-distributed. Sums of non-iid exponential random variables may not be gamma-distributed.

For large numbers of independent loans, we could use the Central Limit Theorem. However, some CDS and CDOpt underliers are large portfolios of related loans and CDO tranches. Exposure to common risk factors can induce correlations in individual loan defaults. Correlated defaults and tranching make these underliers behave like portfolios of fewer loans.

I propose that the average underlier default time — a sum of non-identically-distributed, possibly correlated summands — is often nearly gamma-distributed.

Approximating these sums with a gamma distribution involves finding a similarly-behaving sum of (often fewer) iid exponential random variables.

These iid default times allow us to easily model underlier default times via an iid-equivalent loan portfolio.

The number of loans (summands) in the iid-equivalent loan portfolio yields an iid-equivalent loan count (ILC). The ILC measures the minimal number of identical but unrelated loans needed to create a similarly (default) risky portfolio. Thus the ILC is a measure of loan portfolio diversification.

Approximating correlated defaults has two benefits. First, it improves models of default times for loan portfolios and CDO tranches. This allows us to price CDOpts and might help us price CDSs. Second, it allows us to better assess the default-relative diversification of loan portfolios.

2. Thinking About Default Times

Current thinking on delays can be traced to Erlang's (1909) pioneering paper: Erlang models the answering delay of busy operators as exponentially distributed.

Jarrow and Turnbull (1995) first suggested modeling bond default times as exponentially-distributed. Jarrow, Lando and Turnbull (1997) model the default time of each credit rating with bonds changing credit ratings via a Markov chain.

Banasik, Crook and Thomas (1999) briefly consider default times that are exponentially- or Weibull-distributed. Collin-Dufresne, Goldstein and Hugonnier (2004) model default times as exponentially-distributed with a random intensity and discuss modeling a two-loan CDO. However, they do not generalize their result.

Correlated defaults have received less attention. Jarrow and Yu (2001) use the Jarrow-Turnbull model to study default by issuers with bond crossholdings. However, they only consider two bonds since for more firms "working out these distributions is more difficult."

Since those distributions are difficult to work out, approximation would seem to be in order. My approach is similar in spirit to what Duffie, et al. (2009) call "frailty-correlated" default[1]; however, I allow for asymptotic approximation.

Approximating the distribution of a random variable was first addressed by the small-sample asymptotic work of Thiele (1871), Gram (1883), and Edgeworth (1883) (the Edgeworth and Gram-Charlier series).

[1]Neither Duffie et al. (2009) nor I use true frailty-based defaults. Frailty, as in survival analysis, is theoretically troublesome: such models assign finite probability to the simultaneous default of all loans.

De Forest (1883), Romanovsky (1924), and Wishart (1926) found results increasingly close to those here. Yet none of these use cumulant differences, discuss the order of error, nor mention earlier work on small-sample asymptotics.

Patnaik (1949), Cox and Reid (1987), and McCullagh (1987) discuss similar approximations. Cox and Barndorff-Nielsen (1989) approximates a weighted sum of Exp(1) variables with a gamma base density and hints at expanding upon this[2]. Surprisingly, none of these use approximations of the form here.

3. Approximation Consistency

We might assume that the time to default[3] for one loan is exponentially- or nearly gamma-distributed (if random sub-events alter the default rate). An increase in defaults might follow a rise in interest rates, a tightening of credit standards, or an economic downturn. Thus we might expect default times for multipl loans to be positively correlated. This means that portfolios of a large number of loans might behave like portfolios of fewer loans.

Finding the average default time for a portfolio involves adding multiple exponential random variables. Note that averages can be composed as sums of rate-changed exponentials.

One or more summands (aka constituents) model a borrower repaying a loan. I assume one constituent is idiosyncratic to the borrower; the other constituents are related to risk factors. Loans which share risk factors have correlated defaults. For a loan portfolio, if we can find a default-equivalent portfolio of independent loans, we can then model the distribution of portfolio defaults.

I consider three cases, in increasing generality:

1) known number of constituents, homogeneous rates;
2) known number of constituents, heterogeneous rates; and,
3) unknown number of constituents, rates.

3.1. **Notation and Assumptions.** We begin by letting:

k = the number of risk factors;
m = the number of loans and risk factors (*i.e.* constituents);
X_i = default time of loan $i \in \{1, \ldots, m\}$;
λ_i = the rate parameter characterizing default time X_i; and,
Y = the time to complete portfolio default.

[2]See example 4.4.

[3]Default may be censored: borrowers might be in good-standing at observation time; or, borrowers might pay back a loan at maturity. These difficulties do not challenge the validity of a delay-based model.

I also assume that: (i) sum constituents are exponentially-distributed; and, (ii) all observations of complete portfolio defaults are independent.

3.2. **Known Number of Independent Constituents, Homogeneous Rates.** This case assumes that borrowers default at a common exponential rate $\lambda_i = \lambda$ but that their defaults are independent.

For constituents of equal importance, the distribution of the sum Y is a Gamma(m, λ) random variable. For constituents of unequal importance, let $w_i > 0$ be the weight assigned to X_i. We then rewrite the moment generating function (mgf) using $\lambda_i^* = \lambda_i / w_i$.

If larger loans are more likely to default, importance-based weighting may counter the λ_i variation and drive these sums closer to being gamma-distributed.

Theorem 1. *(When Default Rates Scale with Loan Size)*
If $\forall\ i \in \{1, \ldots, m > 1\}$*:*

1) $X_i \overset{indep}{\sim} \text{Exp}(\lambda_i)$*; and,*
2) *there exists a weight* $0 < w_i < \infty$ *such that* $\lambda_i = w_i \lambda_i^* = w_i \lambda$*,*

then $Y = \sum_{i=1}^{m} w_i X_i \sim \text{Gamma}(m, \lambda)$.

Proof. The mgf exists in a neighborhood about $t = 0$, identifying the distribution. $\mathcal{M}_Y(t) = \prod_{i=1}^{m} \frac{\lambda_i}{\lambda_i - w_i t} = \prod_{i=1}^{m} \frac{\lambda}{\lambda - t}$, which is the mgf for a Gamma(m, λ) random variable. □

If the λ_i^*'s are not equal, we have a known number of heterogeneous rates.

3.3. **Known Number of Constituents, Heterogeneous Rates.** Since the mgf for a sum of independent random variables is the product of the individual mgf's, we get the moment and cumulant generating function (cgf) for this case:

$$\mathcal{M}_Y(t) = \prod_{i=1}^{m} \frac{\lambda_i}{\lambda_i - t}, \tag{1}$$

$$\mathcal{K}_Y(t) = \sum_{i=1}^{m} \left(\log \lambda_i - \log(\lambda_i - t)\right), \tag{2}$$

as well as the first four cumulants of the sum Y:

$$\kappa_1 = \sum_{i=1}^{m} \frac{1}{\lambda_i}, \qquad \kappa_3 = \sum_{i=1}^{m} \frac{2}{\lambda_i^3},$$
$$\kappa_2 = \sum_{i=1}^{m} \frac{1}{\lambda_i^2}, \qquad \kappa_4 = \sum_{i=1}^{m} \frac{6}{\lambda_i^4}.$$

Since the mgf and cgf depend on the individual rates, we must find the density explicitly for each problem instance. This can be cumbersome for many rates but may be handled by the following case.

3.4. **Unknown Number of Constituents, Rates.** Most flexible is to assume nothing about the λ_i's, m nor independence. Instead, we approximate the average default density. Edgeworth's (1883, 1905, 1906) work suggests expanding about a base density to get an approximate density.

Since correlation is an important aspect of what we are studying, we should ensure that Edgeworth expansions are consistent for correlated exponentials. The following proof shows consistency for certain (structural) constructions of correlated exponentials:

Theorem 2. *(Consistency for Exponentials Correlated by Risk Factors) Assume the following hold:*

1) $Y = \sum_{i=1}^{m} X_i$; $X_i \sim \text{Exp}(\lambda_i)\ \forall\ i \in \{1, \ldots, m\}$;
2) *X_i's are partitioned by an independent (idiosyncratic risk factors; $\bar{\mathcal{S}}$) and singular (systematic risk factors; $\mathcal{S}_1, \ldots, \mathcal{S}_k$) index sets;*
3) *at least two of these index sets are non-empty*[4]*;*
4) *X_i's in the independent index set ($X_{i\in\bar{\mathcal{S}}}$'s) are independent; and,*
5) *X_i's belonging to different index sets are independent. (Thus all risk factors and idiosyncratic risks are independent.)*

Then Edgeworth expansions are consistent for estimating the density of Y.

Proof. By assumption 2, we may put $X_{S_\ell} := X_{i\in\mathcal{S}_\ell}$ with rate λ_{S_ℓ} for all $\ell \in \{1, \ldots, k\}$. We then rewrite Y as $Y = \sum_{i\in\bar{\mathcal{S}}} X_i + |\mathcal{S}_1| X_{S_1} + \ldots + |\mathcal{S}_p| X_{S_p}$.

Since $X_{S_1}, \ldots, X_{S_p}$ are exponentially-distributed, we can rewrite $|\mathcal{S}_i| X_{S_i}$ as an exponential random variable $X^*_{S_i}$ with rate $\lambda^*_{S_i} = \lambda_{S_i}/|\mathcal{S}_i|$. Thus we get $Y = \sum_{i\in\bar{\mathcal{S}}} X_i + \sum_{i=1}^{k} X^*_{S_i}$.

This is a sum of independent, non-identically distributed constituents. This meets the regularity conditions in Feller (1971) since Y has finite higher moments and the characteristic function $\phi_Y(t)$ is integrable for $m > 1$. □

[4]Were this not true, a risk factor switching on could cause all loans to immediately default.

4. Approximation Forms

With the preceding consistency result, we may investigate asymptotic approximations. I examine three possibilities:

1) Normal-correction Edgeworth Density Approximation;
2) Gamma-correction Edgeworth Density Approximation; and,
3) Mélange Edgeworth Density Approximation.

4.0.1. *Normal-Correction Edgeworth Density Approximation.* For an Edgeworth approach, we might consider a distribution with support over all of $\mathbb{R}$. This can be done via an expansion about the normal distribution. We choose a normal-based expansion and match the first two cumulants:

$$f_Y(y) = \frac{\phi(z)}{\sqrt{\kappa_2}} \left[1 + \frac{\kappa_3(z^3 - 3z)}{6\sqrt{\kappa_2^3}} + \frac{\kappa_4(z^4 - 6z^2 + 3)}{24\kappa_2^2} + \frac{\kappa_3^2(z^6 - 15z^4 + 45z^2 - 15)}{72\kappa_2^3}\right] + O(n^{-3/2}) \tag{3}$$

where $z = (y - \kappa_1)/\sqrt{\kappa_2}$.

4.0.2. *Gamma-correction Edgeworth Density Approximation.* As noted previously: most Edgeworth expansions use the normal distribution. However, the preceding sections suggest we expand about the gamma distribution.

To clarify the results, I introduce two more pieces of notation:

$\gamma_{m,\lambda}(y) =$ the Gamma(m, λ) pdf if $m > 0$, 0 otherwise; and,
$\Gamma_{m,\lambda}(y) =$ the Gamma(m, λ) cdf if $m > 0$, 0 otherwise.

Next I define $\gamma^{(k)}_{m,\lambda}(y)$ as a bounded differentiation of $\gamma_{m,\lambda}(y)$:

$$\gamma^{(k)}_{m,\lambda}(y) = \lambda^k \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \gamma_{m-j,\lambda}(y) I_{m-k>0}.$$

Thus $\gamma^{(k)}_{m,\lambda}(y) = 0$ for all negative $m - k$. This upholds the regularity condition of a bounded k-th derivative; details are in Feller's (1971) second volume on page 538.

We recall the Gamma(m, λ) cumulants and match the first two, implying

$$\hat{m} = \frac{\kappa_1^2}{\kappa_2}, \qquad \hat{\lambda} = \frac{\kappa_2}{\kappa_1} \tag{4}$$

for estimates of m and λ. This and the preceding derivatives yield:

$$
\begin{aligned}
f_Y(y) = \gamma_{\hat{m},\hat{\lambda}}(y) &+ \frac{\kappa_3\hat{\lambda}^3 - 2\hat{m}}{6}\sum_{j=0}^{3}(-1)^{3-j}\binom{3}{j}\gamma_{\hat{m}-j,\hat{\lambda}}(y) \\
&+ \frac{\kappa_4\hat{\lambda}^4 - 6\hat{m}}{24}\sum_{j=0}^{4}(-1)^{4-j}\binom{4}{j}\gamma_{\hat{m}-j,\hat{\lambda}}(y) \\
&+ \frac{(\kappa_3\hat{\lambda}^3 - 2\hat{m})^2}{72}\sum_{j=0}^{6}(-1)^{6-j}\binom{6}{j}\gamma_{\hat{m}-j,\hat{\lambda}}(y) \\
&+ O(n^{-3/2}),
\end{aligned} \tag{5}
$$

assuming that $\hat{m} \geq 7$ to meet the aforementioned regularity condition.

Note that the expansion has a pleasingly simple form: binomial sums of other gamma densities. These approximations are elegant and, so far as I can tell, a new result.

4.0.3. *Mélange Edgeworth Density Approximation.* We might consider a mélange of the preceding approaches[5]: a base distribution close to the true distribution coupled with simple correction terms. While unusual, the idea is recommended by Cox and Barndorff-Nielsen (1989).

A sensible mélange here is to use a Gamma(m, λ) base density and add normal-correction terms. This eliminates concerns about correction terms not existing due to $\hat{m}$ being too small.

$$
f_Y(y) = \gamma_{\hat{m},\hat{\lambda}}(y) + \frac{\phi(z)}{\sqrt{\kappa_2}}\left[\mathcal{C}_3(z,\kappa) + \mathcal{C}_4(z,\kappa)\right] + O(n^{-3/2}) \tag{6}
$$

where, as before, $z = (y - \kappa_1)/\sqrt{\kappa_2}$ with

$$
\begin{aligned}
\mathcal{C}_3(z,\kappa) &= \frac{\kappa_3 - 2\hat{m}/\hat{\lambda}^3}{6\sqrt{\kappa_2^3}}(z^3 - 3z); \quad \text{and,} \\
\mathcal{C}_4(z,\kappa) &= \frac{\kappa_4 - 6\hat{m}/\hat{\lambda}^4}{24\kappa_2^2}(z^4 - 6z^2 + 3) \\
&\quad + \frac{(\kappa_3 - 2\hat{m}/\hat{\lambda}^3)^2}{72\kappa_2^3}(z^6 - 15z^4 + 45z^2 - 15).
\end{aligned}
$$

[5]The term "mélange" is used to avoid confusion with mixture distributions.

5. Evaluating the Approximations

To examine how these approximations perform, I examine the following (tractable) forms for approximating the default time density:

1) Normal base and normal Edgeworth terms;
2) Gamma base density;
3) Mélange: Gamma base and normal Edgeworth terms; and,
4) Gamma base and gamma Edgeworth terms.

These approximations are compared to densities for default times of large CDO tranches. The CDO setup is consistent with information in Lucas (2001) and Fender and Kiff (2004):

1) underlying portfolio of 200 equal-sized loans;
2) each loan has a different rate of default;
3) four tranches, with defaults allocated to the lowest still-extant tranche:

Tranche	# Loans	Percent
A	150	75%
Mezzanine	40	20%
Equity	10	5%.

The equity tranche for this CDO is likely the most difficult tranche to model. Therefore, I focus on approximating the equity tranche under the possibility of a major shock. This shock induces correlations via large (500%) default acceleration.

200,000 simulated loan portfolios were created using the algorithm in Appendix A. These simulations yielded cumulants which implied parameters for the Edgeworth approximations. The target density was then plotted along with the approximations. (((TO DO: CALCULATE GOODNESS-OF-FIT METRICS!)))

5.1. CDO Equity Tranche with Major Shock.

The average default time of a CDO equity tranche can be modeled as a mean of correlated exponential random variables. Correlations are induced by a shock (*e.g.* a precipitate economic downturn). In particular, suppose the following event setup:

1) the correlation-inducing systematic component is a rare one-time economic shock;
2) each loan i not in default ignores the shock with probability p_i; and,

$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_4$
0.142	2.583×10^{-3}	1.018×10^{-4}	6.458×10^{-6}

Table 1. Simulated cumulants for weighted average default time of a CDO equity tranche modeled as the mean of the smallest ten of 200 correlated non-identically-distributed exponential random variables with five-fold acceleration of default after an exogenous shock.

3) shock-affected loans have their remaining time to default accelerated by a factor[6] δ.

Finally, we make some modeling assumptions:

1) default rates are chosen to cover a range of average (idiosyncratic, ex-shock) default times of 5–20 years:
$$\{\lambda\}_i = \left\{ \frac{10^{i/200}}{20} : i \in 1, \ldots, 200 \right\};$$
2) the systematic shock has rate $\lambda_s = 0.05$ (mean time-to-shock of 20 years) which induces theoretical pre-reaction correlations[7] from 0.048 (λ_i's near 0.5) to 0.330 (λ_i's near 0.05);
3) the probability of ignoring a shock is $p_i = 0$;
4) the default time acceleration factor $\delta = 5$.

For example, suppose in simulation that loan A defaults in 21 years and loan B in 3 years. If the shock happens in year 10, loan A would default in year $10 + (21 - 10)/5 = 12.2$ and loan B would still default in year 3.

We compute the average default time as the sample mean ($\bar{Y}$) of the ten smallest random variables. Simulated $\bar{Y}$'s yielded the sample cumulants in Table 5.1 and implied gamma parameters of $\hat{\lambda} = 55.12$ and $\hat{m} = 7.849$.

The average simulated default time is about two months; and, the iid-equivalent loan count is 7.8 — a reduction in default-relative diversification of nearly one-quarter. These are stark indicators of the default risk taken on by equity tranche holders.

Plots of the normal-correction approximations (Figure 1) to the equity tranche average default time density show a few key details:

- the standard (normal) approximation performs very well with minor negativity for $\bar{y} < 0.04$;

[6]By the memoryless property of the exponential distribution, each shock-affected loan defaults at $X_s + X_i'$ where $X_s \sim \text{Exp}(\lambda_s)$ and $X_i \sim \text{Exp}(\delta\lambda_i)$. Thus Theorem 2 still applies.

[7]cum-reaction correlations are lower

- the mélange expansion has almost no negativity and is indistinguishable from the actual density more often than the standard Edgeworth approximation.

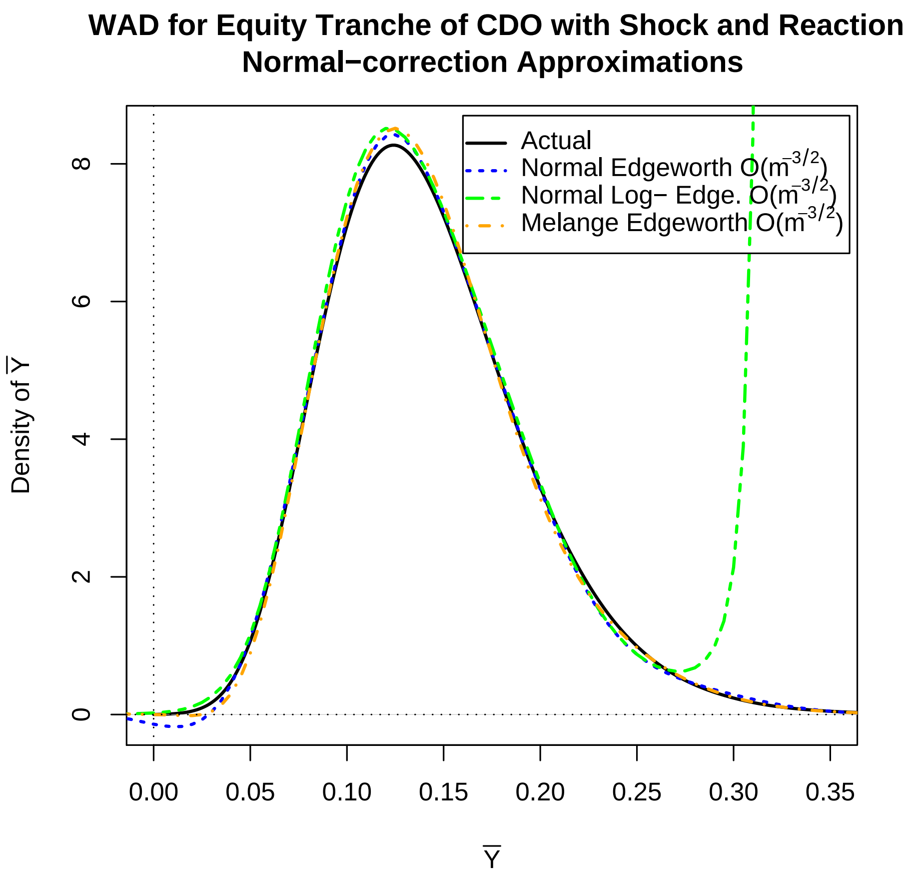

FIGURE 1. CDO equity tranche weighted average default time (WAD) density (solid line) versus approximations. $O(m^{-3/2})$ normal approximation (dotted line); $O(m^{-3/2})$ normal-correction log-density approximation (dash-dot-dotted line); and, the $O(m^{-3/2})$ mélange approximation (dash-dotted line). The equity tranche is modeled by the smallest ten of 200 correlated exponential random variables with five-fold default acceleration after an exogenous shock.

Plots of the gamma-correction approximations (Figure 2) to the equity tranche average default time density show that:

- the gamma base is almost identical to the actual density; and,

- the gamma-correction approximation is positive for $0 < \bar{y} < 0.05$).

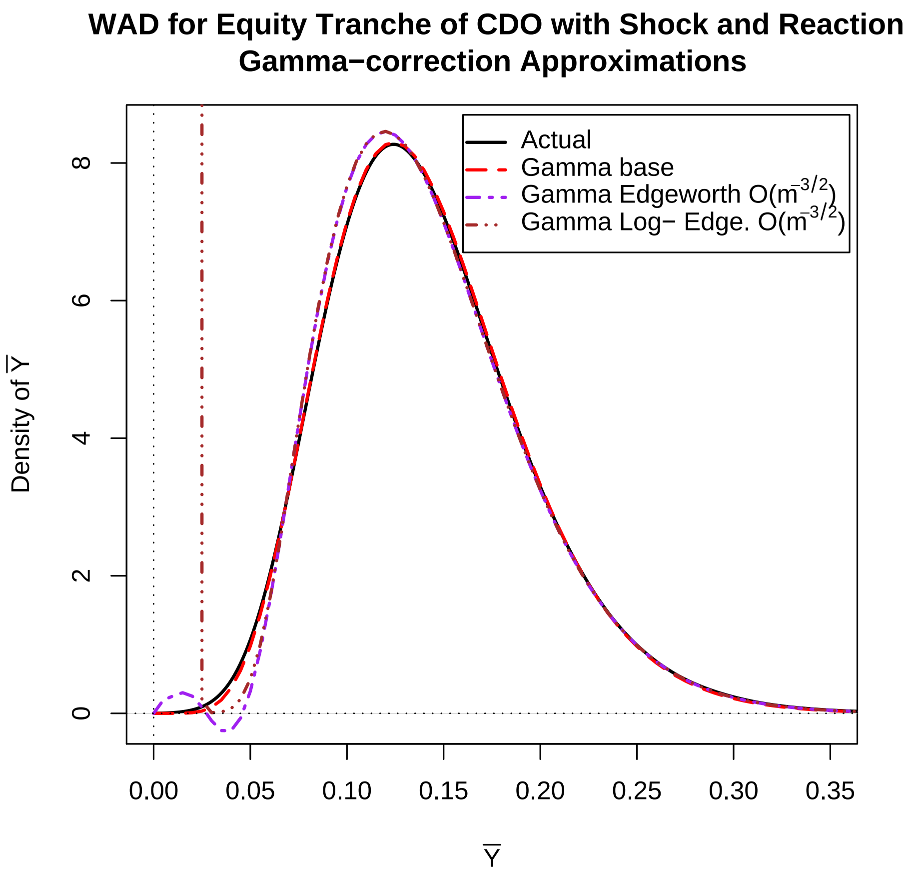

FIGURE 2. CDO equity tranche weighted average default time (WAD) density (solid line) versus approximations. $O(m^{-1/2})$ gamma base (– –); $O(m^{-3/2})$ gamma-correction approximation (— - -); and, the $O(m^{-3/2})$ gamma-correction log-density approximation (−⋯). The equity tranche is modeled by the smallest ten of 200 correlated exponential random variables with five-fold default acceleration after an exogenous shock.

5.2. **Other CDO Tranches with Major Shock.** These plots naturally raise some questions:

1) What do average default time densities look like for other tranches?
2) What if $p_i = 0.5$ (or some other number)?

3) What if p_i is inversely-proportional to λ_i? (*i.e.* What if issuers with better credit are less likely to face default acceleration?)

For $p_i = 0$, the other tranche average default time densities exhibit left-skew which increases with tranche seniority. The A tranche average default time, in particular, is sharply left-skewed with a minor mode on the left, a non-zero plateau in the middle, and a major mode on the right.

If the probability of ignoring the shock increases to $p_i = 0.5$ or $p_i = 0.8$, the A tranche average default time looks normally-distributed while the mezzanine tranche exhibits barely-discernible right-skew. The equity tranche average default time has slightly less right-skew than for $p_i = 0$; but, even a casual observer would probably doubt normality.

Finally, if the probability of ignoring the shock is proportional to credit quality, the A and mezzaanine tranche average default times appear normally-distributed; the equity tranche average default time density is right-skewed (as for $p_i = 0$).

6. Conclusion and Future Research

As the examples make clear, the average default time for a loan portfolio may be nearly-gamma-distributed. In many cases the distribution is approximated well by a gamma-based Edgeworth approximation.

The gamma base alone is attractive for its parsimony and other nice behavior: guaranteed positivity, unimodality, and computational ease. In some cases, the gamma base plus gamma-correction terms nicely approximate the distribution in question.

The gamma base and gamma-correction Edgeworth expansions also have another advantage: their tail decay is on the order of e^{-y} instead of e^{-y^2}. While not detectable from the plots, this is an important difference for analyses involving extreme events: the standard Edgeworth approximations would predict far fewer extreme events.

In some situations, the mélange approximation may be useful because it offers quickly-decaying correction terms while still offering tail decay on the order of e^{-y}.

The greatest need for further work is to investigate if distributions of average default times help in modeling distributions of portfolio default times. One simple possibility would be to use the average default time distribution along with the ILC. The portfolio would then be modeled as experiencing total default after all $\hat{m}$ iid-equivalent loans had defaulted.

One disappointing feature of gamma-correction Edgeworth expansions not shown here is their poor performance at approximating distributions which are left-skewed.

An effective way to handle this might be to choose a value M and model the left-skewed distribution with y-reversed gamma base or correction terms originating from $y = M$. The reversed gamma densities used would be of the form $\gamma_{m,\lambda}(M - y)$. A mixture of standard and reversed gamma densities could be even be used, dictated by the signs of the cumulant differences.

Another area for further work is to study when the implied gamma base parameter $\hat{m}$ is close to violating regularity conditions. In these cases, it might be fruitful to bias $\hat{m}$ upward so that gamma-correction terms may be used. How this would affect overall performance is unclear.

While standard Edgeworth procedure is to match the first two moments, this may not be optimal. One could investigate the performance of Edgeworth expansions where the pseudocumulants are determined by maximum likelihood or by minimizing some measure of the distance between the approximate and actual densities.

The performance of such maximum-likelihood "Edgeworth expansions" is surely better than using pseudocumulants; however, the approximation order is then a model selection question. Such an approach would probably incorporate higher-order cumulant effects via the optimization — and thus might be between the Edgeworth and saddlepoint expansions in spirit.

Appendix A. Generating Positively Correlated Defaults

This method uses the memoryless property of the exponential distribution to efficiently reuse idiosyncratic random variates.

Algorithm 1. *Positively-correlated (nearly-gamma) default times*

1) *Generate idiosyncratic components* $\tilde{X}_i \overset{indep}{\sim} \text{Exp}(\beta_i)$.
2) *Generate systematic components* $\tilde{X}_{s,k} \sim \text{Exp}(\beta_{s,k})$.
3) *Sort the systematic shocks to find which occur first:* $X_{s,(1)} < \cdots < X_{s,(K)}$.
4) *Reorder the acceleration coefficients* δ_k*'s similarly.*
5) *For* $k = 1$ *to* K
   - **(a)** *Find affected loans:* $\mathcal{J}_k = \{i : \text{loan } i \text{ exposed to risk } k\} \cap \{i : \tilde{X}_i > \tilde{X}_{s,(k)}\}$.
   - **(b)** *Accelerate defaults in affected loans:* $\tilde{X}_{i \in \mathcal{J}_k} = X_{s,(k)} + (\tilde{X}_{i \in \mathcal{J}_k} - X_{s,(k)})/\delta_{(k)}$.

References

Banasik, J., J. N. Crook, and L. C. Thomas (1999): Not If But When Will Borrowers Default. *J. Oper. Res. Soc.* 50(12), 1185–1190.

Collin-Dufresne, P., R. Goldstein, and J. Hugonnier (2004): A General Formula for Valuing Defaultable Securities. *Econometrica* 72(5), 1377–1407.

Cox, D. R. and O. E. Barndorff-Nielsen (1989): *Asymptotic Techniques for Use in Statistics*. London: Chapman and Hall.

Cox, D. R. and N. Reid (1987): Approximations to Noncentral Distributions. *Can. J. Stat./Rev. Can. Stat.* 15(2), 105–114.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009): Frailty Correlated Default. *J. Finan.* 64(?), ?–?

Edgeworth, F. Y. (1883): On the Method of Ascertaining a Change in the Value of Gold. *J. Stat. Soc. London* 46(4), 714–718.

Edgeworth, F. Y. (1905): The Law of Error. *Trans. Cambr. Phil. Soc.* 20, 35–65,113–141.

Edgeworth, F. Y. (1906): The Generalised Law of Error, or Law of Great Numbers. *J. Roy. Stat. Soc.* 69(3), 497–539.

Erlang, A. K. (1909): The Theory of Probabilities and Telephone Conversations. *Nyt Tids. Math.* B(20), 33–39.

Feller, W. (1971): *An Introduction to Probability Theory and Its Applications*, vol. II. 2nd edn. New York: John Wiley and Sons.

Fender, I. and J. Kiff (2004): CDO Rating Methodology: Some Thoughts on Model Risk and its Implications. Working Paper 163, Bank for International Settlements.

Forest, E. L. D. (1883): On an Unsymmetrical Probability Curve [Continued]. *Analyst* 10(1), 1–7.

Gluck, J. and H. Remeza (2000): Moody's Approach to Rating Multi-Sector CDOs. Structured Finance Special Report (September), Moody's Investors Service.

Gram, J. P. (1883): Über die Entwickelung reeler Funktionen in Reihen mittelstder Methode der kleinsten Quadrate. *J. reine angew. Math.* 94, 41–73.

Jarrow, R. A., D. Lando, and S. M. Turnbull (1997): A Markov Model for the Term Structure of Credit Risk Spreads. *Rev. Finan. Stud.* 10(2), 481–523.

Jarrow, R. A. and S. M. Turnbull (1995): Pricing Derivatives on Financial Securities Subject to Credit Risk. *J. Finan.* 50(1), 53–85.

Jarrow, R. A. and F. Yu (2001): Counterparty Risk and the Pricing of Defaultable Securities. *J. Finan.* 56(5), 1765–1799.

Lucas, D. (2001): *CDO Handbook*. New York: JP Morgan Securities.

McCullagh, P. (1987): *Tensor Methods in Statistics*. London: Chapman and Hall.

Patnaik, P. B. (1949): The Non-Central χ^2- and F-Distribution and their Applications. *Biometrika* 36(1/2), 202–232.

Romanovsky, V. (1924): Generalisation of Some Types of the Frequency Curves of Professor Pearson. *Biometrika* 16(1/2), 106–117.

Severini, T. (2005): *Elements of Distribution Theory.* Cambridge University Press.

Thiele, T. N. (1871): *En mathematisk Formel for Dødeligheden, prøvet paa en af Livsforsikringanstalten af 1871 benyttet Erfaringrække.* Copenhagen: C. A. Reitzel.

Thiele, T. N. (1872): On a Mathematical Formula to Express the Rate of Mortality Throughout the Whole of Life. *J. Instit. Act.* 16, 313–329. Translated by T. B. Sprague.

Wishart, J. (1926): On Romanovsky's Generalised Frequency Curves. *Biometrika* 18(1/2), 221–228.