# Approximating correlated defaults

Rosenthal, Dale W.R.

University of Illinois at Chicago

2008

# APPROXIMATING CORRELATED DEFAULTS

DALE W.R. ROSENTHAL

Abstract. Modeling defaults is critical to risk management as well as pricing debt portfolios and portfolio derivatives. In the recent financial crisis, multi-billion-dollar losses resulted from correlated defaults that were improperly modeled. This paper proposes statistical approximations which are more general than those used previously, follow from an intensity-based risk-factor model, and allow consistent parameter estimation. The parameters imply an approximating portfolio of independent, identical-credit loans and characterize both average credit quality and default-relative diversification (aka the "diversity score"). Unlike previous approaches, these metrics are derived jointly from theory. The approach addresses weaknesses in the typical diversity score-based methods by allowing for fatter tails as well as loans differing in size and credit quality. The approximations may also be used to model complete portfolio default and help set capital adequacy requirements. An example shows how to estimate the approximating portfolio.

(*JEL*: G13, G12, C16, G33)

In the financial crisis of 2008–2009, US households lost about \$11 trillion in wealth and structured debt securities alone suffered impairments of over \$1.5 trillion.[1] Many loan portfolios experienced high numbers of correlated defaults: defaults at an accelerated rate and clustered in time. This

---

Helpful comments were provided by Per Mykland, Michael Wichura, and participants at the SoFiE Geneva, Peking Guanghua QMBA, and Oakland credit analysis conferences. Financial support from the National Science Foundation under grants DMS 06-04758 and SES 06-31605 is gratefully acknowledged.

Address correspondence to Dale W.R. Rosenthal, Department of Finance, University of Illinois at Chicago, 601 S. Morgan St (MC 168), Chicago, IL 60607; e-mail: daler@uic.edu.

[1]Statistics on the scope of the crisis and the effect on structured products are well documented in Financial Crisis Inquiry Commission (2011) and Tung *et al.* (2011).

correlation was greater than predicted by typical methods such as Schorin and Weinreich (1998). One of the causes and magnifiers of the crisis noted by the Financial Crisis Inquiry Commission was the increased correlation of risky assets. In particular, the US Federal Reserve's Troubled Asset Relief Program and Term Asset-Backed Securities Loan Facility programs were targeted at assets which had been affected by a larger-than-expected number of concurrent defaults.

Modeling correlated defaults well is of particular importance because of the real effects of defaults. Many of these effects were seen in the recent financial crisis. Financial firms withdrew funding liquidity from world markets; financial and non-financial firms failed due to the loss of funding; and, unemployment increased while tax revenues decreased. Many investors not directly affected by these troubles were affected by the increased market volatility. Some investors rebalanced their portfolios at a time when doing so was costly. Other investors chose not to fund risky ventures which further suppressed job growth.

Since defaults are so destructive, lenders often manage default risk by creating portfolios of loans. These portfolios may be tranched or split into portfolio credit derivatives such as collateralized debt obligations (CDOs). Default risks of portfolios, tranches, or other derivatives may also be hedged with credit default swaps (CDSs). A primary concern for any portfolio, however, is risk correlations. Two approaches dominate the modeling of default: structural models of firm assets and liabilities and reduced-form models of default intensities.

This paper shows how the assumptions behind some reduced-form approaches may be too restrictive. We propose approaches that relax these assumptions to find default-approximating portfolios. We then show how these portfolios may characterize the diversification and aggregate credit quality of a loan portfolio. The source of these two measures is more robust than many reduced-form methods because the assumptions are weaker. The measures also better capture the tail risk of default times.

To get at the correlation of such defaults, we propose modeling the time to loan default as an idiosyncratic component which interacts with shared systematic components (risk factors). Since shared risk factors lead to default correlation, we explore approximating expansions for the distribution of correlated default times. The expansions imply approximating portfolios of independent, identical-credit loans.

The parameters of these approximations have direct economic meaning for the approximating portfolios: they measure both default-relative diversification of a correlated loan portfolio and aggregate credit quality. The default-relative diversification is measured via an iid-equivalent loan count and yields a theoretically-derived maximum likelihood version of Moody's KMV's diversity score.[2] One contribution of this paper is to derive the diversity score jointly with the average credit quality so that the two are consistent. Furthermore, one of the approximating expansions has a form which is mathematically concise and theoretically novel.

## 1. Thinking About Default Times

Since we can think of the time to loan default as random, it makes sense to ask when such a default is more likely to occur. The structural modeling approach assumes that default happens at a random time determined by the stochastic evolution of a firm's assets (and perhaps liabilities) with

[2]An explanation of a common approach to calculating the Moody's KMV diversity score is given in Schorin and Weinreich (1998) and recapped in Duffie and Gârleanu (2001). Duffie and Gârleanu (2001) also has a theory-based derivation of the diversity score although they do not jointly estimate the average credit quality.

default occuring when some barrier is hit by the asset process. The reduced-form intensity-based approach does not model the firm and instead assumes defaults of individual loans happen stochastically at some rate.

If we observe a firm's assets and the liability-dependent default barrier, we may model default via structural models such as in Merton (1974), Black and Cox (1976), and Leland and Toft (1996). Were we to observe these data for multiple firms, we could use asset correlations with a default barrier to study default correlation as in Zhou (2001). Giesecke (2006), however, noted that we often cannot directly observe firm assets and that this leads us to use an intensity-based, reduced-form modeling approach.

Reduced-form models can be traced back to theory in Erlang (1909) which modeled the answering delay of busy operators as exponentially distributed. Jarrow and Turnbull (1995) first suggested the modeling of bond default times as exponentially distributed. Jarrow *et al.* (1997) modeled default times of many credit ratings with Markov switching of bonds among credit ratings; and, Banasik *et al.* (1999) briefly considered exponential or Weibull default times. Collin-Dufresne *et al.* (2004) modeled default times as exponential with random intensities and discuss modeling a two-loan CDO.

Modeling correlated defaults is more complex and thus came later. Jarrow and Yu (2001) used the Jarrow-Turnbull model to study default by two firms with bond cross-holdings; and, they note that for more firms "working out these distributions is more difficult." This complexity explains the popularity of reduced-form approaches. Duffie and Gârleanu (2001) decomposed a firm's (exponential) default rate into systematic and idiosyncratic components, an approach also used in this paper. Giesecke (2003) considered exponentially distributed defaults correlated by a joint distribution with exponential marginals and a singular "spine."[3] Duffie *et al.* (2009) analyzed which random variables affect the default intensity process using a linear model.

The approximations used here follow from theory in Edgeworth (1883). Patnaik (1949), Cox and Reid (1987), and McCullagh (1987) discussed similar approximations; and, Cox and Barndorff-Nielsen (1989) approximated a weighted sum of Exp(1) variables with a gamma density and hinted at a possible expansion. However, none of those approaches use approximations of the form here.

## 2. Reduced-Form Model and Approximation Consistency

We start with a reduced-form model that assumes the time to default for one loan is exponentially or nearly gamma-distributed.[4] The gamma distribution has not generally been used in the reduced-form default-intensity literature. However, if other random events alter the default rate, the actual distribution of defaults may well be gamma-distributed (as is shown later). Furthermore, the gamma is a less restrictive distributional assumption since the exponential distribution is a special case of the gamma (with 1 degree of freedom).

We assume multiple possible events affect the rate at which a borrower defaults. One event is idiosyncratic to the borrower; others are related to risk factors. These risk events interact with the idiosyncratic event. For example, defaults may increase with rising interest rates, tightening credit standards, or an economic downturn. The possibility that loans may share risk factors

[3]This construction is the same as the failure-time distribution first studied by Marshall and Olkin (1967). Simultaneous defaults are theoretically problematic since they imply a singular component of the joint distribution which breaks the assumptions of most modeling approaches. For this reason, we do not follow this approach.

[4]Default may be censored, including by loan maturation. This does not challenge the validity of a delay-based model.

yields positively-correlated default times for multiple loans. Portfolios of many loans would then behave like portfolios of fewer loans.

The assumption of risk factors affecting default intensities is not new; others have used affine models to explore shared risk factors. While affine models are simpler, they can be problematic if used to model bonds of differing credits. Affine models imply that bonds of differing credit experience the same additive change in their default rate with respect to a risk factor. For example, a AAA-rated bond in recession would default at a rate $\lambda_{\mathrm{AAA}} + \gamma$ and a B-rated bond would default at a rate $\lambda_{\mathrm{B}} + \gamma$ where $\lambda_{\mathrm{B}} > \lambda_{\mathrm{AAA}}$. Instead, we assume that the default rate changes by a multiplicative factor (*e.g.* $\delta\lambda_{\mathrm{AAA}}$ and $\delta\lambda_{\mathrm{B}}$) with respect to a risk factor. Further support for using a non-affine model comes from Das *et al.* (2007) who found clustering beyond that suggested by an affine model.

Unfortunately, working with a non-affine model is more difficult than working with an affine model. However, if we can find a default-approximating portfolio of independent identical-credit loans, we can then model the distribution of portfolio defaults. To help find this portfolio, we focus on the distribution of the average time to default.[5] While we would prefer to work with the default time distribution, working with the distribution of the average allows us to prove asymptotic consistency. This approach was also used by Schorin and Weinreich (1998).

2.1. **Reduced-Form Model.** We begin by setting up notation. Let:

$k$ = number of risk factors/possible risk events;

$\ell$ = number of loans/possible idiosyncratic default events;

$X_i$ = delay before an event i, $i \in \{1, \ldots, \ell + k\}$;

[5] Finding the average default time for a portfolio involves adding multiple exponential random variables (the idiosyncratic default times and the shared-risk-factor event times). We also make use of the fact that averages can be written as sums of rate-changed exponentials.

$\lambda_i$ = rate parameter characterizing delay $X_i$;

$\delta_j$ = rate multiplier of $\lambda_{i\in\{1,\ldots,\ell\}}$ after risk event $j \in \{1,\ldots,k\}$;

$Y_i$ = time to default for loan $i \in \{1,\ldots,\ell\}$; and,

$\mathcal{F}_t$ = filtration encapsulating information known up to time $t$.

The model assumes event times are exponentially distributed. For each loan $i \in \{1,\ldots,\ell\}$, the rate parameter $\lambda_i$ implies an idiosyncratic probability of default in a given year and thus a certain credit quality. For each risk factor $j \in \{1,\ldots,k\}$, the rate parameter $\lambda_{\ell+j}$ implies a probability of a risk event occurring in a given year and $X_{\ell+j}$ is the time of that occurrence.

To model default times, we assume a relationship of events for each loan. First, loans default at their idiosyncratic default rates $\lambda_{i\in\{1,\ldots,\ell\}}$. Second, some loans may be exposed to a risk factor $j$ and undefaulted when a risk factor event occurs. The idiosyncratic default rate of these loans accelerates to $\delta_j\lambda_i$ after that risk event. For example, homeowners might default at a certain rate but at twice that rate after the economy enters a recession.

2.2. **Homogenous, Independent Loans.** We start by considering loans with no risk factors and only idiosyncratic propensities to default. This approach helps develop the case where loans have shared risk factors.

For independent loans of equal credit quality ($\lambda_{i\in\{1,\ldots,\ell\}} = \lambda$), the average default time $\bar{Y} = \frac{1}{\ell}\sum_{i=1}^{\ell} X_i$ is a Gamma$(\ell, \ell\lambda)$ random variable. For loans of unequal size, let $w_i > 0$ be the portfolio weight of loan $i$. If larger loans are more likely to default, notional weighting may drive the average default time to being gamma-distributed.

**Theorem 1.** *(When Default Rates Scale with Loan Size) If $\ell > 1$:*

1) $X_i \overset{indep}{\sim} \mathrm{Exp}(\lambda_i) \quad \forall\ i \in \{1,\ldots,\ell > 1\}$*; and,*

2) *there exist weights $0 < w_i < \infty$ such that $\lambda_i/w_i = \ell\lambda$,*

*then $\bar{Y} = \sum_{i=1}^{\ell} w_i X_i \sim \mathrm{Gamma}(\ell, \ell\lambda)$.*

*Proof.* The mgf exists in a neighborhood about $t = 0$ and is integrable for $\ell > 1$, identifying the distribution. $\mathcal{M}_{\bar{Y}}(t) = \prod_{i=1}^{\ell} \frac{\lambda_i}{\lambda_i - w_i t} = \prod_{i=1}^{\ell} \frac{\ell\lambda}{\ell\lambda - t}$, which is the mgf for a Gamma$(\ell, \ell\lambda)$ random variable. □

If the $\lambda_i / w_i$ quotients are not equal, we essentially have heterogeneous rates. Since the mgf for a sum of independent random variables is the product of the individual mgf's, we get the mgf and cumulant generating function (cgf):

$$\mathcal{M}_{\bar{Y}}(t) = \prod_{i=1}^{\ell} \frac{\lambda_i}{\lambda_i - w_i t}, \tag{1}$$

$$\mathcal{K}_{\bar{Y}}(t) = \sum_{i=1}^{\ell} \left(\log \lambda_i - \log(\lambda_i - w_i t)\right), \tag{2}$$

and the first four cumulants of $\bar{Y}$: $(\kappa_1, \kappa_2, \kappa_3, \kappa_4) = \sum_{i=1}^{\ell} \left(\frac{w_i}{\lambda_i}, \frac{w_i^2}{\lambda_i^2}, \frac{2w_i^3}{\lambda_i^3}, \frac{6w_i^4}{\lambda_i^4}\right)$.

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 portfolios of many loans.

2.3. **Reduced Form Consistency.** We may progress further if we assume nothing about independence, the $\lambda_i$'s, nor the number of risk factors $k$. Instead, we approximate the density of the average default time $\bar{Y}$. Edgeworth (1883, 1905, 1906) suggests expanding about a base density to get an approximate density.

To more easily express event times, we let $X_{(j)}$ be the order statistics for the risk event times $X_{\ell+1}, \ldots, X_{\ell+k}$ such that $X_{(1)} \leq \cdots \leq X_{(k)}$. We also let $\delta_{(j)}$ be the rate acceleration parameter for $X_{(j)}$ and set $k_i^*$ to be the time of the most recent risk event which affected loan $i$:

$$k_i^* = \arg\max_j X_{(j)} I(X_{(j)} \leq t) I(i \in \mathcal{S}_j). \tag{3}$$

By the memoryless property of the exponential distribution, the time $Y_i$ to default on loan $i \in \{1, \dots, \ell\}$ may be written as a sum of a rate-changed idiosyncratic $X_i$ and systematic $X_{j \in \{\ell+1, \dots, \ell+k\}}$'s. This allows us to express the default time $Y_i$ of loan $i$ as

$$Y_i|\mathcal{F}_t = \sum_{j=1}^{k_i^*} (X_{(j)} - X_{(j-1)}) + \frac{X_i - X_{(k_i^*)}}{\prod_{j=1}^{k_i^*} \delta_{(j)}} = \frac{X_i - X_{(k_i^*)}}{\prod_{j=1}^{k_i^*} \delta_{(j)}} + X_{(k_i^*)} \tag{4}$$

$$\stackrel{\mathcal{L}}{=} \frac{X_i}{\prod_{j=1}^{k_i^*} \delta_{(j)}} + X_{(k_i^*)}. \tag{5}$$

Simulation of these random variables is explained further in Section 4.1.

To approximate correlated loan portfolios, we must ensure Edgeworth expansions may be used. The following proof shows consistency for the mean of the previously-outlined construction of default times:

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

1) *$X_i$'s are partitioned by index sets: one independent $\bar{\mathcal{S}}$ (idiosyncratic risk factors) and many singular $\mathcal{S}_1, \dots, \mathcal{S}_k$ (systematic risk factors).*
2) *If loan $i \in \{1, \dots, \ell\}$ is exposed to risk factor $j \in \{1, \dots, k\}$, then $i \in \mathcal{S}_j$.*
3) *At least two of these index sets are non-empty.*[6]
4) *$X_i$'s in the independent index set ($X_{i \in \bar{\mathcal{S}}}$'s) are independent.*
5) *$X_i$'s belonging to different index sets are independent. (Thus all risk factors and idiosyncratic risks are independent.)*
6) $\bar{Y} = \frac{1}{\ell} \sum_{i=1}^{\ell} Y_i$.
7) $X_i \sim \mathrm{Exp}(\lambda_i)\ \forall\ i \in \{1, \dots, \ell + k\}$.

[6]Were this not true, a risk event could cause all loans to immediately default. That would yield a non-zero probability of $X_i = X_{i'}$ for $i \neq i'$. Such a singularity would prevent us from using these approximations. This means that the techniques used here might fail for the Giesecke (2003) model. However, at a fine enough granularity of time simultaneous default becomes less likely.

*Then Edgeworth expansions are consistent estimators of the $\bar{Y}$ density.*

*Proof.* By assumption 1, we may put $X_{S_j} := X_{i\in\mathcal{S}_j}$ with rate $\lambda_{S_j}$ for all $j \in \{1, \ldots, k\}$. We then rewrite $\bar{Y}$ as

$$\bar{Y}|\mathcal{F}_T = \frac{1}{\ell}\left(\sum_{i\in\bar{\mathcal{S}}} \frac{X_i}{\prod_{j=1}^{k} \delta_j^{I(i\in\mathcal{S}_j)}} + |\mathcal{S}_1|X_{S_1} + \ldots + |\mathcal{S}_k|X_{S_k}\right) \tag{6}$$

$$= \frac{1}{\ell}\left(\sum_{i\in\bar{\mathcal{S}}} X_i^* + |\mathcal{S}_1|X_{S_1} + \ldots + |\mathcal{S}_k|X_{S_k}\right), \tag{7}$$

where $X_i^*$ is a rate-changed exponential random variable with rate $\lambda_i \prod_{j=1}^{k} \delta_j^{I(i\in\mathcal{S}_j)}$.

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 $\bar{Y} = \sum_{i\in\bar{\mathcal{S}}} X_i^* + \sum_{i=1}^{k} X_{S_i}^*$.

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

**Corollary 1.** *(Individual Default Times Are Nearly Gamma-Distributed) Under the assumptions above, if a loan is exposed to at least one systematic risk factor, the default time for that single loan is nearly gamma-distributed (and possibly not exponentially distributed).*

*Proof.* Given the proceeding proof, we need only note that:

$$Y_i = \frac{X_i}{\prod_{j=1}^{k} \delta_j^{I(i\in\mathcal{S}_j)}} + \sum_{j=1}^{k} I(i \in \mathcal{S}_j) X_{S_j} \tag{8}$$

If there is at least one systematic risk factor to which loan $i$ is exposed, then the sum meets the regularity conditions outlined above and the unconditional distribution is not exponential. □

## 3. Approximation Forms

With consistency assured, we investigate three asymptotic approximations: (1) normal Edgeworth expansions; (2) gamma Edgeworth expansions; and, (3) a combination of the above ("mélange").[7]

### 3.1. Normal Edgeworth Expansions.

Most Edgeworth expansions use a normal base density and expand about the normal density yielding multiplicative correction terms involving Hermite polynomials. This creates a distribution approximation with support over all of $\mathbb{R}$. We begin with such an expansion by matching the first two cumulants:

$$
\begin{aligned}
f_{\bar{Y}}(y) = \frac{\phi(z)}{\sqrt{\kappa_2}} \Bigg[ 1 &+ \frac{\kappa_3(z^3 - 3z)}{6\sqrt{\kappa_2^3}} + \frac{\kappa_4(z^4 - 6z^2 + 3)}{24\kappa_2^2} \\
&+ \frac{\kappa_3^2(z^6 - 15z^4 + 45z^2 - 15)}{72\kappa_2^3} \Bigg] + O(\ell^{-3/2})
\end{aligned}
\tag{9}
$$

where $z = (y - \kappa_1)/\sqrt{\kappa_2}$. While the correction terms allow us to match skewness and kurtosis, the tail decay is of order $e^{-z^2}$ which may be problematic as discussed in Duffie and Gârleanu (2001).

### 3.2. Gamma Edgeworth Expansions.

While most Edgeworth expansions are based on the normal density, Section 2 suggests the default time density might be close to a gamma density. If we use a gamma base density and expand about that, the resulting correction terms are binomial sums of other gamma densities differing only in the degrees of freedom. This approach also has the structural advantage of assigning zero probability to negative default times.

To clarify the results, we let:

[7] Approximation of the log-density, was also explored. Such forms are non-negative but may behave poorly in the tails. For the example in Section 4.2, log-density approximations were superior for the middle of the distribution but exploded in the right and left (near 0) tails. Because of this instability, log-density expansions are omitted.

$$\begin{aligned}\gamma_{\ell,\lambda}(y) &= \text{the Gamma}(\ell,\lambda) \text{ pdf if } \ell > 0, \text{else } 0; \text{ and,}\\ \gamma^{(k)}_{\ell,\lambda}(y) &= \text{the } k\text{-th bounded derivative of } \gamma_{\ell,\lambda}(y), \text{else } 0\\ &= \lambda^k \sum_{j=0}^{k} (-1)^{k-j} \binom{k}{j} \gamma_{\ell-j,\lambda}(y) I_{\ell-k>0}.\end{aligned}$$

Thus $\gamma^{(k)}_{\ell,\lambda}(y) = 0$ for all negative $\ell - k$. This upholds the regularity condition of a bounded $k$-th derivative as in Feller (1971), page 538.

We recall the Gamma$(\ell, \lambda)$ cumulants and match the first two, implying $\hat{\ell} = \kappa_1^2/\kappa_2$ and $\hat{\lambda} = \kappa_2/\kappa_1$ for estimates of the effective $\ell$ and $\lambda$. This and the preceding derivatives yield:

$$\begin{aligned}f_{\bar{Y}}(y) = \gamma_{\hat{\ell},\hat{\lambda}}(y) &+ \frac{\kappa_3\hat{\lambda}^3 - 2\hat{\ell}}{6} \sum_{j=0}^{3} (-1)^{3-j} \binom{3}{j} \gamma_{\hat{\ell}-j,\hat{\lambda}}(y)\\ &+ \frac{\kappa_4\hat{\lambda}^4 - 6\hat{\ell}}{24} \sum_{j=0}^{4} (-1)^{4-j} \binom{4}{j} \gamma_{\hat{\ell}-j,\hat{\lambda}}(y)\\ &+ \frac{(\kappa_3\hat{\lambda}^3 - 2\hat{\ell})^2}{72} \sum_{j=0}^{6} (-1)^{6-j} \binom{6}{j} \gamma_{\hat{\ell}-j,\hat{\lambda}}(y) + O(\ell^{-3/2}),\end{aligned} \tag{10}$$

assuming that $\hat{\ell} \geq 7$ to meet the aforementioned regularity condition.[8]

Note that the expansion has a pleasingly simple form: binomial sums of other gamma densities. That the expansion takes this form is an elegant and novel result. We know of no other Edgeworth expansion (apart from the typical normal-based form) that has such a concise and easily-expressed form. Since we have shown that default times are likely to be well approximated by a gamma distribution, this expansion also gives us the machinery to find that approximation.

[8]We can relax this to just $\hat{\ell} \geq 4$ with an error bound of only $O(\ell^{-1})$ if we drop the last two binomial sums.

The correction terms allow us to match skewness and kurtosis (as in the normal expansion). The tails are fatter with decay of order $e^{-z}$. These features help handle concentration risk and reduce the underestimation of tail risk discussed in Duffie and Gârleanu (2001) and problematic in the Schorin and Weinreich (1998) method, for example. For risk managers, this model more effectively captures tail risk.

3.3. **A Combination of Expansions (Mélange).** We might instead combine the preceding approaches: a base density thought to be close to the true density, but coupled with correction terms from the normal density. This approach allows for mathematially tractable correction terms and is recommended by Cox and Barndorff-Nielsen (1989). Since the term "combination" is too prevalent to refer to this approach alone and "mixture" is used in other contexts, we dub this approach a "mélange" for clarity and convenience.

One possible mélange uses a Gamma$(\ell, \lambda)$ base density and normal-derived correction terms. This eliminates concerns about correction terms not satisfying regularity conditions if $\hat{\ell}$ is too small; however, it assigns some small probability to negative times.

$$f_{\bar{Y}}(y) = \gamma_{\hat{\ell},\hat{\lambda}}(y) + \frac{\phi(z)}{\sqrt{\kappa_2}}\left[\mathcal{C}_3(z,\kappa) + \mathcal{C}_4(z,\kappa)\right] + O(\ell^{-3/2}) \tag{11}$$

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

$$\mathcal{C}_3(z,\kappa) = \frac{\kappa_3 - 2\hat{\ell}/\hat{\lambda}^3}{6\sqrt{\kappa_2^3}}(z^3 - 3z); \quad \text{and,}$$

$$\mathcal{C}_4(z,\kappa) = \frac{\kappa_4 - 6\hat{\ell}/\hat{\lambda}^4}{24\kappa_2^2}(z^4 - 6z^2 + 3) + \frac{(\kappa_3 - 2\hat{\ell}/\hat{\lambda}^3)^2}{72\kappa_2^3}(z^6 - 15z^4 + 45z^2 - 15).$$

This expansion also has fatter tails ($e^{-z}$ decay, if the correction terms are used) like the preceding gamma-based expansion.

## 4. Evaluating the Approximations

Having developed the theory, we can explore how well these expansions approximate the average default time density for CDO tranches. Since gamma Edgeworth regularity conditions may not hold or the expansion may be negative, we also examine the base gamma density with no correction terms.

The CDO setup is consistent with information in Lucas (2001) and Fender and Kiff (2004): an underlying portfolio of 200 equal-sized loans. Each loan has a different rate of default, chosen to equally cover a range of log-default rates. The CDO trust is divided into three tranches, with defaults allocated to the lowest still-extant tranche (Table 1).

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

Table 1. Tranche structure for a 200-loan CDO with three tranches. The percentages are in keeping with typical values discussed in Lucas (2001) and Fender and Kiff (2004).

While this CDO setup differs from later standardization on 125-loan CDOs, the results here would be even stronger for a 125-loan CDO. For a 125-loan CDO, limiting asymptotic approximations would be less accurate and thus would make these (large deviation) expansions perform relatively better.

Default correlations are induced by the possibility of one (*i.e.* $k = 1$) rare systematic risk event such as a sharp economic downturn at time $X_{\ell+1}$. Each loan $i$ not in default experiences accelerated default: the remaining time to default is divided by a factor $\delta_1$. (By the memoryless property, Theorem 2 still applies.[9])

We assume a range of equally spaced log-default rates. Since the log-default rates are equally spaced, this setup is more difficult to handle than if that range supported a modal distribution of log-default rates. These yield average (idiosyncratic, unshocked) default times of 5–20 years with a skew toward better credits: $\lambda_i = 10^{i/200}/20$ for $i \in 1, \ldots, 200$. The systematic shock occurs at a rate $\lambda_{\ell+1} = 0.05$, implying a mean time-to-shock of 20 years.[10] Since the one-time shock is serious, the default time acceleration used is $\delta_1 = 5$. These are equivalent to a portfolio of B–C-rated loans which become C–D-rated during a recession.

4.1. **Generating Risk Factor-Correlated Defaults.** 200,000 simulated loan portfolios were created using Algorithm 1. The simulations yielded default times for each loan. Cumulants of those default times were calculated, which implied parameters for the expansions. The target density was then

---

[9]By memorylessness, shocked loans default at $X_{\ell+1} + X_i'$ for a risk event at time $X_{\ell+1} \sim \text{Exp}(\lambda_{\ell+1})$, $X_i' \sim \text{Exp}(\delta\lambda_i)$.

[10]If all live loans defaulted at shock time, this would induce correlations of 0.048 ($\lambda_i$'s near 0.5) to 0.330 ($\lambda_i$'s near 0.05) as per Marshall and Olkin (1967). Since the default merely accelerates, we expect correlations less than these values.

plotted along with the approximations, and goodness-of-fit measures were calculated.

The simulation method uses the memoryless property of the exponential distribution. Specifically, we use the fact that the distribution of the time remaining to default is conditionally independent of a stopping time. This property implies that the time remaining has the same distribution as the original default time (given the information at time 0), and allows us to efficiently reuse idiosyncratic random variates.

**Algorithm 1.** *(Risk Factor-Correlated Default Times)*

1) *Generate idiosyncratic default times:* $\tilde{X}_{i\in\{1,\ldots,\ell\}} \overset{indep}{\sim} \mathrm{Exp}(\lambda_i)$.
2) *Generate systematic shock time:* $\tilde{X}_{\ell+1} \sim \mathrm{Exp}(\lambda_{\ell+j})$.
3) *For* $i = 1$ *to* $\ell$*, process each loan.*
   **(a)** *If* $\tilde{X}_i > \tilde{X}_{\ell+1}$*:*
   *Accelerate remaining default time in affected loans.*
   **1:** *Set* $\tilde{X}_i := (\tilde{X}_i - \tilde{X}_{\ell+1})/\delta_1$.
   **2:** *Combine time periods:*
   $X_i = \tilde{X}_i + \tilde{X}_{\ell+1}$.
   **(b)** *Otherwise:* $X_i = \tilde{X}_i$.

Appendix A explains how to do this sort of a simulation in a setting with multiple risk factors, which is a bit more involved.

4.2. **Example: 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. The equity tranche of a CDO is the riskiest and most likely to be affected by correlated defaults. Thus we focus on approximating the equity tranche under the possibility of a major shock, which would induce correlations via default acceleration.

The average default time is the sample mean ($\bar{Y}$) of the ten smallest generated default times. Simulated $\bar{Y}$'s yielded the sample cumulants of $(\hat{\kappa}_1, \hat{\kappa}_2, \hat{\kappa}_3, \hat{\kappa}_4) = (0.142, 2.583 \times 10^{-3}, 1.018 \times 10^{-4}, 6.458 \times 10^{-6})$ which implied gamma parameters of $\hat{\lambda} = 55.12$ and $\hat{\ell} = 7.849$.

The average simulated default time is about two months; and, the estimated diversity score (iid loan count) is 7.8. This implies that the shared systematic risk factor causes a reduction in default-relative diversification of nearly one-quarter. Further, we can think of the equity tranche as having default behavior like a portfolio of 7.8 loans each with an annual probability of default of $1 - e^{-55.12} \approx 1$ (F credit). These are stark indicators of the default risk taken by equity tranche holders in this setup.

Plots of the approximations (Figure 1) to the average default time density show overall good performance. The normal Edgeworth expansion is slightly negative for $\bar{y} < 0.04$; the gamma Edgeworth expansion is negative for $0.025 < \bar{y} < 0.05$; and, the gamma base is almost identical to the actual density. The mélange (gamma base, normal Edgeworth correction terms) has almost no negativity and is close to the actual density more often than the standard normal Edgeworth approximation.

While the gamma Edgeworth approximation has better tail behavior, it is unfortunately not the best approximation. The mean squared errors (Table 2) suggest that the simple gamma base, which still yields economically meaningful metrics, or the mélange approach are better approximations.

| Normal Edgeworth | Gamma Base | Gamma Edgeworth | Mèlange |
|---|---|---|---|
| 0.0034 | 0.0006 | 0.0306 | 0.0051 |

Table 2. Mean-squared errors for density approximations. Note that the $O(\ell^{-1/2})$ gamma base is more accurate than the other, $O(\ell^{-3/2})$ approximations.

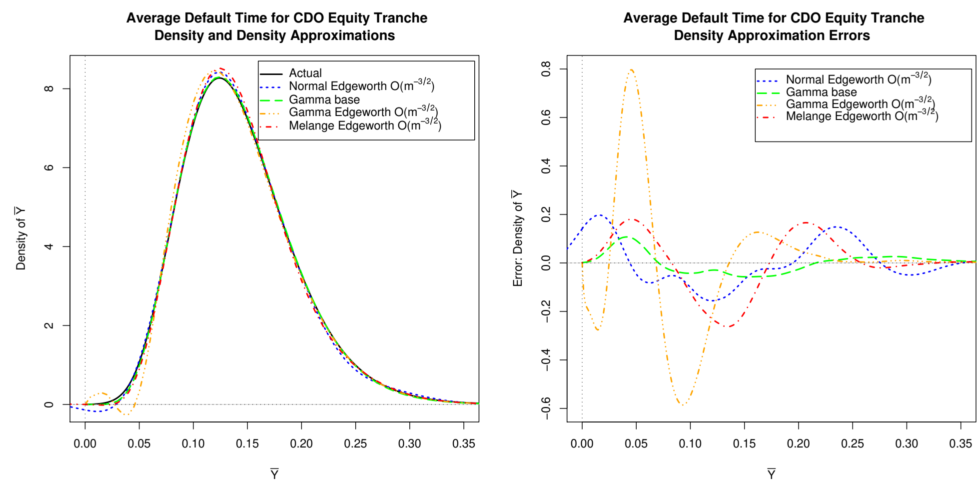


FIGURE 1. CDO equity tranche average default time density approximations (left) and errors (right). The true density is the solid line. Approximations: normal Edgeworth (- -); gamma base (— —); gamma Edgeworth ($-\cdots$); and, gamma-normal mélange ($-\cdot$). All are $O(\ell^{-3/2})$ except the $O(\ell^{-1/2})$ gamma base.

4.3. **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 loans ignored risk events with some probability $p_i > 0$?
3) What if $p_i$ is inversely-proportional to $\lambda_i$? (*i.e.* What if issuers with better credit are less likely to face default acceleration?)

For $p_i = 0$ (as above), average default time densities for other tranches exhibit left-skew which increases with tranche seniority. The A tranche density, 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 a risk event is higher ($p_i = 0.5, 0.8$), 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 mezzanine tranche average default times appear normally-distributed. The equity tranche average default time density is right-skewed (as for $p_i = 0$).

## 5. Estimating the Approximating Portfolio

We estimate the approximating portfolio in three steps. First, we estimate the arrival rate for a systematic risk event (*e.g.* crisis). Next, we estimate the idiosyncratic default rates and their acceleration during a systematic risk event. Finally, we use these parameters to simulate loan portfolios from which we measure cumulants. The cumulants specify the approximating portfolio.

To estimate the occurence rate for the systematic risk event, the maximum-likelihood estimator is simply the number of events per unit of time. For example, if our systematic risk event were a US recession, we could use the NBER business cycle data to estimate $\lambda_s$. Since the NBER notes an average cycle length of 55 months, we would estimate $\lambda_s = 0.218$.

To estimate the default acceleration $\delta$ during a risk event, we use data about loans similar to those in our portfolio and which existed when a systematic risk event occurred at some time $t_s$. If we characterize the loan credit quality for one of those similar loans $i$ as $q(i)$ (*e.g.* $q(i) \in \{$good, mediocre, poor$\}$), we can find the default acceleration parameter by maximizing the likelihood function given default times $t_i$:

$$\mathcal{L}(\lambda, \delta | t_s) = \underbrace{\prod_{i \in \{\text{defaulted}\}, t_i < t_s} \lambda_{q(i)} e^{-\lambda_{q(i)} t_i} \cdot (1 - e^{-\lambda_s t_s})}_{\text{pre-crisis defaults}} \times$$
$$\underbrace{\prod_{i \in \{\text{defaulted}, t_i \geq t_s\}} \delta \lambda_{q(i)} e^{-\delta \lambda_{q(i)} (t_i - t_s)} \cdot e^{-\lambda_s t_s}}_{\text{in-crisis defaults}} \times$$
$$\underbrace{\prod_{i \in \{\text{undefaulted,active}\}} e^{-\delta \lambda_{q(i)} (t - t_s)} \cdot e^{-\lambda_s t_s}}_{\text{possible future defaults}}.$$
$$\underbrace{\prod_{i \in \{\text{undefaulted,repaid}\}} e^{-\delta \lambda_{q(i)} (T_i - t_s)} \cdot e^{-\lambda_s t_s}}_{\text{undefaulted (censored default)}}. \quad (12)$$

Ideally, the idiosyncratic default rates $\lambda_{q(i)}$ would be the same as those in the portfolio we are concerned with approximating. However, even if this is not the case, we can still get a reasonable estimate of $\delta$ if we are willing to assume a constant distress acceleration of default. We then use the estimated $\delta$ to find the approximating portfolio through simulation. This yields the metrics of interest: the diversity score and the average credit quality.

5.1. **Example: 25-loan Subprime Portfolio.** As an example of such inference, we analyze a portfolio of 25 C-credit loans ($q(i) = \text{poor}$).

We are immediately faced with two possible approaches: We could use risk-neutral default rates for pricing or physical default rates for risk management (and an appropriate stochastic discount factor for pricing).

To get risk-neutral default rates, we could look at credit default swaps (CDSs) for the bonds in our portfolio and assume these did not price in default correlation with risk factors. Similarly, we could find government, agency, and municipal bonds that track our risk factors and look at CDSs on those bonds to imply a risk-neutral arrival rate for those risks. Unfortunately, the assumption that bond CDSs do not price in the possibility of correlated defaults is unlikely since CDSs are forward-looking instruments.

Instead, we look at physical default rates. This has the advantage of cleanly separating the effect of idiosyncratic defaults from risk factors; and, this is the measure we would want for risk management.

We start by examining (*ex post*) a portfolio of similar (C-credit) loans from another recession to find the typical level of default acceleration $\delta$ and idiosyncratic default rate $\lambda_{q(i)}$. For a twenty-loan portfolio with the default times given in Table 3, we would maximize the likelihood equation (12) to estimate a crisis default-acceleration factor of $\hat{\delta} = 3.28$ and an idiosyncratic default rate of $\hat{\lambda}_{\text{poor}} = 0.22$.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Pre-crash defaults | 3.2 | 4.8 | 5.7 | | | | | | | |
| Post-crash defaults | 5.8 | 5.8 | 5.8 | 5.9 | 5.9 | 6.0 | 6.1 | 6.2 | 6.3 | 6.5 |
| | 6.8 | 7.2 | 7.7 | 8.3 | 9.2 | | | | | |
| Repaid | 10.0 | 10.0 | | | | | | | | |

TABLE 3. Table of loan lifetimes with 3 defaults before a market crash at 5.8 years; 15 defaults after the crash; and, 2 undefaulted loans which were repaid in full at ten years. The repaid loans are "right censored" since we do not observe default and inference must account for this phenomenon.

Returning to the original portfolio of 25 C-credit loans, we can conduct an *ex ante* analysis using the estimated $\hat{\delta}$ and $\hat{\lambda}_{\text{poor}}$.[11] At time $t = 0$, we assume that $\lambda_{\text{poor}} = 0.22$ (as estimated).

The interaction between the systematic risk factor and the loans is not affine, so we cannot easily find a closed-form solution to generate cumulants. Instead, we proceed via simulation as outlined in Algorithm 1. This simulation is straightforward since we have only one systematic risk factor.

[11] This entails the assumption that the estimated crisis default-acceleration rate $\hat{\delta}$ is valid for future crises.

Because the inference of $\hat{\lambda}$ and $\hat{\ell}$ assumes we have good estimates of cumulants $\hat{\kappa}_1$ and $\hat{\kappa}_2$, we average the mean and variance of default times.[12] For 10,000 simulations, this gives us the cumulants in Table 4. These cumulants, mean $\hat{\kappa}_1$ and variance $\hat{\kappa}_2$, jointly determine the iid-loan credit quality $\hat{\lambda} = \frac{\hat{\kappa}_2}{\hat{\kappa}_1}$ and the diversity score (number of iid loans) $\hat{\ell} = 25\widehat{\mathrm{df}} = 25\frac{\hat{\kappa}_1^2}{\hat{\kappa}_2}$ for the approximating portfolio.

| $\hat{\kappa}_1$ | $\hat{\kappa}_2$ | $\hat{\lambda}$ | $\widehat{\mathrm{df}}$ | $\hat{\ell}$ |
|---|---|---|---|---|
| 3.170 | 6.368 | 0.498 | 0.634 | 15.841 |

Table 4. Table of simulation-averaged cumulants for 10,000 simulations of 25 individual C-credit (subprime) loans with default acceleration of $\delta = 3.28$ in a recession arriving at NBER-implied rates. The shared systematic risk factor induces default correlation which, in turn, affects the cumulants. The cumulants imply the default intensity $\hat{\lambda}$ and diversity score $\hat{\ell}$ for a 25-loan subprime portfolio. Note that the default-equivalent portfolio is one of just under 16 D-credit loans.

From this analysis, we see that our portfolio of 25 C-credit loans has default behavior best approximated by a portfolio of just under 16 D-credit loans. This represents a 37% reduction in diversification strictly due to the default acceleration from a shared macroeconomic risk factor. Clearly, the capital adequacy required to hold such a portfolio would be significantly greater than that required for holding 25 independent C-credit loans unaffected by other risk factors.

## 6. Conclusion and Future Research

As the above analyses show, the expansions we propose can help analyze the effects of defaults correlated by default-accelerating risk factors. The

[12] In this section, we switch from working with the average default time in a portfolio to working with individual loan default times. While the asymptotic proofs required working with average default times, the inference here does not require that — as shown in Chambers (1967). Since our approximation is for iid loans, the only correction needed is to then multiply the implied degrees of freedom by 25.

greatest value from these approximations is the idea of an approximating portfolio of independent, identical-credit loans and the two theoretically-based, consistent metrics that approximation yields: $\hat{\ell}$, the diversity score (number of iid loans in the approximating portfolio), and $\hat{\lambda}$, the credit quality of those loans. While Duffie and Gârleanu (2001) estimated the diversity score alone, they note that they do not estimate the average loan credit quality. Thus this paper shows both a different way to estimate the diversity score and jointly estimates the diverstiy score with the average loan credit quality.

A further benefit is that many of these approximations perform well at approximating the average default density. Ideally, we want the distribution of total default; and, these approximations may be useful to achieving that. One possibility would be to use the average default time distribution along with the iid-equivalent loan count. The portfolio would then be modeled as experiencing total default after all $\hat{\ell}$ iid loans had defaulted.

As for the approximations themselves, the gamma Edgeworth expansion is mathematically concise and novel. Furthermore, the gamma Edgeworth expansion, the mélange, and even the base gamma density all have tail behavior that is thicker than standard Edgeworth expansions: tail decay on the order of $e^{-y}$ instead of $e^{-y^2}$. Furthermore, the gamma and mélange expansions add correction terms to better capture the tail risk which Duffie and Gârleanu (2001) raised as an issue with most reduced-form approaches. While not detectable from the plots, these features are important for analyses involving extreme events: the standard Edgeworth approximations (and even most reduced-form approaches) would predict far fewer extreme events.

One shortcoming of gamma Edgeworth expansions not shown here is their poor performance at approximating distributions which are left-skewed.[13] An effective way to handle this might be to use the maturity time $T$ to model the left-skewed distribution with a $y$-reversed gamma base or correction terms originating from $y = T$. The reversed gamma densities used would be of the form $\gamma_{\ell,\lambda}(T - y)$. 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 $\hat{\ell}$ is close to violating regularity conditions. In these cases, is may be fruitful to bias $\hat{\ell}$ upward so the gamma-correction terms may be used.

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 both accuracy and spirit.

## Appendix A. Multi-Risk Factor-Correlated Default Times

To simulate default times affected by multiple shocks, we first must keep track of which loans are affected by which shocks. Therefore, we set:

$$\mathcal{A}_i := \{j \in \{1, \ldots, \ell\} : \text{loan } i \text{ affected by risk factor } j\}. \tag{A.1}$$

[13] Simulations of CDO A tranches suggested that their average default times were left-skewed. Therefore, the ideas mentioned here are of more than idle concern.

**Algorithm 2.** *(Multi-Risk Factor-Correlated Default Times)*

1) *Generate idiosyncratic default times:* $\tilde{X}_{i\in\{1,\dots,\ell\}} \overset{indep}{\sim} \mathrm{Exp}(\lambda_i)$.
2) *Generate systematic shock times:* $\tilde{X}_{j\in\{\ell+1,\dots,\ell+k\}} \sim \mathrm{Exp}(\lambda_{\ell+j})$.
3) *Sort the systematic shock times* $X_{\ell+1},\dots,X_{\ell+k}$ *to generate risk-factor order statistics:* $X_{(1)} < \cdots < X_{(k)}$.
4) *Reorder the acceleration coefficients* $\delta_j$ *similarly to get* $\delta_{(j)}$.
5) *For* $i = 1$ *to* $\ell$, *process each loan.*
   **(a)** *Hold last risk factor shock time:* $L_i = 0$.
   **(b)** *For* $j = 1$ *to* $k$, *examine each risk factor (in time order).*
      **1:** *If* $\tilde{X}_i > \tilde{X}_{(j)}$ *and* $(j) \in \mathcal{A}_i$*:*
      *Shock* $(j)$ *affected loan* $i$.
         (A) *Accelerate default due to risk factor* $(j)$ *shock:*
         $\tilde{X}_i = \tilde{X}_i/\delta_{(j)}$.
         (B) *Remember the latest risk shock time:*
         $L_i = \tilde{X}_{(j)}$.
   **(c)** *Set risk factor-affected default time:* $X_i := \tilde{X}_i + L_i$.

## References

Banasik, J., J. N. Crook, and L. C. Thomas (1999): Not If But When Will Borrowers Default. *Journal of the Operational Research Society* 50(12), 1185–1190.

Black, F. and J. C. Cox (1976): Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *Journal of Finance* 31(2), 351–367.

Chambers, J. M. (1967): On Methods of Asymptotic Approximation for Multivariate Distributions. *Biometrika* 54(3/4), 367–383.

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. *Canadian Journal of Statistics / Revue Canadienne de Statistique* 15(2), 105–114.

Das, S. R., D. Duffie, N. Kapadia, and L. Saita (2007): Common Failings: How Corporate Defaults Are Correlated. *Journal of Finance* 62(1), 93–117.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009): Frailty Correlated Default. *Journal of Finance* 64(5), 2089–2123.

Duffie, D. and N. Gârleanu (2001): Risk and Valuation of Collateralized Debt Obligations. *Financial Analyst's Journal* 57(1), 41–59.

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

Edgeworth, F. Y. (1905): The Law of Error. *Transactions of the Cambridge Philosophical Society* 20, 35–65,113–141.

Edgeworth, F. Y. (1906): The Generalised Law of Error, or Law of Great Numbers. *Journal of the Royal Statistical Society* 69(3), 497–539.

Erlang, A. K. (1909): The Theory of Probabilities and Telephone Conversations. *Nyt Tidsskrift for Matematik* 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.

Financial Crisis Inquiry Commission (2011): Final Report of the National Commission on the Causes of the Financial and Economic Crisis in the United States. The Financial Crisis Inquiry Report, United States of America.

Giesecke, K. (2003): A Simple Exponential Model for Dependent Defaults. *Journal of Fixed Income* 13(3), 74–83.

Giesecke, K. (2006): Default and Information. *Journal of Economic Dynamics and Control* 30(11), 2281–2303.

Gram, J. P. (1883): Über die Entwickelung reeler Funktionen in Reihen mittelstder Methode der kleinsten Quadrate. *Journal für die reine und angewandte Mathematik* 94, 41–73.

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

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

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

Leland, H. E. and K. B. Toft (1996): Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. *Journal of Finance* 51(3), 987–1019.

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

Marshall, A. W. and I. Olkin (1967): A Multivariate Exponential Distribution. *Journal of the American Statistical Association* 62(317), 30–44.

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

Merton, R. (1974): On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* 29(2), 449–470.

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

SCHORIN, C. N. and S. WEINREICH (1998): Collateralized Debt Obligation Handbook. Working paper, Fixed Income Research, Morgan Stanley.

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. *Journal of the Institute of Actuaries* 16, 313–329. Translated by T. B. Sprague.

TUNG, J., A. METZ, and N. WEILL (2011): Default and Loss Rates of Structured Finance Securities: 1993–2010. Special comment, Moody's Investors Service.

ZHOU, C. (2001): An Analysis of Default Correlations and Multiple Defaults. *Review of Financial Studies* 14(2), 555–576.