Moody's Correlated Binomial Default Distribution

AUTHOR:

Gary Witt
Managing Director
(212) 553-4352
Gary.Witt@moodys.com

CONTACTS:

Yuri Yoshizawa
Managing Director
(212) 553-1939
Yuri.Yoshizawa@moodys.com

William May
Managing Director
(212) 553-3869
William.May@moodys.com

Gus Harris
Managing Director
(212) 553-1473
Gus.Harris@moodys.com

Investor Liason

Brett Hemmerling
Investor Liaison
(212) 553-4796
Brett.Hemmerling @moodys.com

WEBSITE
www.moodys.com

CONTENTS

1. Motivation: The role of correlation in default distributions
2. Definition of the Correlated Binomial Distribution
   i) Underlying Assumptions
   ii) Closed Form Distribution Formula
3. Parameter Estimation: Choosing the correlation and the correlated diversity score
4. Conclusion
5. Appendices
   a) Appendix I: Comparison of the BET, the Correlated Binomial and simulations using a normal distribution
   b) Appendix II: Derivation of the Correlated Binomial distribution
   c) Appendix III: Computational Issues

SUMMARY

This paper describes Moody's Correlated Binomial default probability distribution. The correlated binomial is similar to Moody's BET but differs from it by explicitly incorporating default correlation. The default correlation is introduced by the assumption that the conditional correlation is constant as defaults increase.

Moody's Investors Service

August 10, 2004

Motivation

In modeling credit risk with Moody's Binomial or BET[1] approach, the actual portfolio of assets underlying a CDO is represented by a reduced number of identical, independent assets. The diversity score, defined as the number of independent assets in the Binomial portfolio, is smaller than the number of assets in the actual portfolio. The purpose of this reduction in the number of assets is to account for the fact that the actual assets are correlated while the Binomial assets are uncorrelated.

This paper describes an alternative way to represent the actual assets that allows the representative assets to be correlated. In the Correlated Binomial[2] approach, the actual portfolio of assets underlying a CDO is represented by a portfolio of identical, correlated assets. Dropping the independence assumption does complicate the calculation of the default probabilities, but it can lead to distributions that have larger probabilities of multiple defaults, sometimes referred to as fat-tailed distributions[3]. In order to recreate the fat-tail effect, the BET is used in conjunction with default stresses that vary with the target rating of the liability. With the Correlated Binomial, the stress factors are unnecessary because correlations are modeled in the representative assets.

Thus the primary motivation behind the introduction of the Correlated Binomial is that the actual assets in a CDO portfolio are correlated and their default distribution has correspondingly higher probabilities of multiple defaults. This feature of the actual default distribution can be more closely modeled by representative assets that are themselves correlated. Not surprisingly, the importance of explicitly modeling the correlation increases as the correlation itself increases. Consequently the Correlated Binomial is currently being used to model cash-flow CDOs backed by portfolios of highly correlated assets that have low diversity scores[4] such as insurance TRUPs and ABS CDOs highly concentrated in one asset type.

Definition

Assume an idealized portfolio of n assets that have identically distributed default distributions[5] with these two properties.

Assumption (1): Each asset has default probability p.

Assumption (2): Each pair of assets has default correlation ρ between them.

Let $x_1, .., x_n$ be random indicator variables representing the default behavior of the assets where $x_j=1$ indicates the default of asset j. Define p_j as the probability of default of asset j given that assets 1 to j-1 are known to have defaulted. Formally, this is written as

$$p_j = E(x_j \mid x_1 = 1, x_2 = 1, ..., x_{j-1} = 1) \quad for\ j = 1,..,n$$

Assumptions (1) and (2) imply that $p_1 = p$ and that $p_2 = p + (1-p)\rho$.

The definition of p_j implies that the probability that k out of k assets default is

$$\prod_{j=1}^{k} p_j = E\left(\prod_{j=1}^{k} x_j\right) \quad for\ k = 1,..,n\,.$$

When $x_1, .., x_n$ are independently distributed, $\rho = 0$ and these assumptions lead to the Binomial distribution with $p_j = p$, a constant for all j = 1, .. ,n. If $\rho > 0$, then $x_1, .., x_n$ are not independent and this is not a Binomial distribution. Furthermore, if $\rho > 0$ and $n > 2$, then assumptions (1) and (2) are not sufficient to determine the joint probability distribution of $x_1, .., x_n$.

1 See "The Binomial Expansion Technique Applied to CBO/CLO Analysis" Moody's *Special Report*, December 1996.
2 The phrase "the Correlated Binomial" is used for convenience. Actually the distribution described in this paper is only one possible way to correlate Bernoulli trials. See Appendix III for a discussion of this point.
3 The fat tails are achieved via the additional parameter, the correlation. All else being equal, increasing the correlation increases the tail probabilities.
4 A diversity score below ten would usually be considered low but other factors such as the average correlation will be considered.
5 Assets that have identically distributed default distributions are sometimes referred to as uniform assets.

Constant Conditional Correlation

In order to specify the joint probability distribution of $x_1, .., x_n$, the Correlated Binomial relies on a third assumption.

Assumption (3): The default correlation between asset j+1 and asset j+2 remains equal to ρ regardless of the number of known defaults among the other j assets.

Put another way, the conditional correlation between any two assets is constant as the number of defaults increases. Mathematically, assumption (3) can be written

Expression (3a): $\rho = Corr(x_{j+1}, x_{j+2} \mid x_1 = 1, x_2 = 1, ..., x_j = 1)$

or equivalently that

Expression (3b): $p_{j+1} = p_j + (1 - p_j)\rho \quad for\ j = 1,.., n-1$

Assumption (3) implies that in the Correlated Binomial p_{j+1}, the default probability of asset j+1 conditional on j defaults, is increasing as j increases. This increasing default probability given other defaults, is one aspect of the fatter tails of the Correlated Binomial distribution[6]. Contrast assumption (3) with the Binomial distribution where the independence assumption implies that $p_j = p$ for all j=1,..,n assets[7].

Closed Form Distribution of Correlated Binomial

Then for k > 0, the probability of k defaults and n-k survivals in any order (analogous to the binomial probabilities used in the BET) is[8]

$$C(n,k)E\left[\prod_{j=1}^{k} x_j \prod_{j=k+1}^{n}(1-x_j)\right] = C(n,k)\sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\prod_{i=1}^{j+k} p_i\right]$$

and the probability of no defaults and n survivals is

$$E\left[\prod_{j=1}^{n}(1-x_j)\right] = 1 + \sum_{j=1}^{n}(-1)^j C(n,j)\prod_{i=1}^{j} p_i$$

Here C(,) represents the combinatorial function $C(n,m) = \dfrac{n!}{m!(n-m)!}$.

The values for p_j can most easily be computed sequentially from Expression (3b) or, if preferred, in closed form equivalent of Expression (3b).

$$p_j = 1 - (1-p)(1-\rho)^{j-1} \quad j = 2,..,n.$$

The larger computational challenge occurs as the number of correlated assets increases and the numerical accuracy of the closed form solution becomes an issue for some software applications[9]. Because of this numerical issue, Moody's will provide on request an Excel spreadsheet with the Correlated Binomial probabili-

6 Another aspect of the fatter tails of the Correlated Binomial is that the probability of survival of asset j+1 given that j assets have survived also increases with j for most values of ρ and p.

7 Since the Correlated Binomial assumes that the correlation between asset j+1 and asset j+2, conditional on the default of the first j assets, is a constant ρ which does not change as j increases, the BET can be seen as a special case of the Correlated Binomial with ρ = 0. This relationship between the Correlated Binomial and the BET is demonstrated in Appendix II.

8 See Appendix II for a derivation of this formula.

9 In Excel 2000, when the number of assets is below fifty there is generally not any numerical instability unless the default probability and correlation are large. The presence of a numerical problem is usually obvious as it results in some probabilities becoming negative.

ties calculated using enhanced arithmetic precision. Also, *Appendix III* discusses this computational issue and contains a high precision algorithm for calculating the Correlated Binomial probabilities.

For a representative portfolio with ten assets and a default probability of 5%, *Table 1* below compares the Binomial to the Correlated Binomial[10]. As the correlation increases, the standard deviation increases but the tail probabilities increase even more dramatically. The last row of *Table 1* shows that as the correlation increases from 0% to 5%, the probability that more than five of the assets default grows from 0.000% to 0.054%.

Table 1

Correlation =	0.00%	2.50%	5.00%
Defaults	**Probabilities**		
0	59.87%	63.07%	65.70%
1	31.51%	26.74%	23.19%
2	7.46%	7.86%	7.74%
3	1.05%	1.88%	2.42%
4	0.10%	0.38%	0.70%
5	0.006%	0.065%	0.185%
6	0.000%	0.009%	0.043%
7	0.000%	0.001%	0.009%
8	0.000%	0.000%	0.001%
9	0.000%	0.000%	0.000%
10	0.000%	0.000%	0.000%
Expected Default Pct	5.00%	5.00%	5.00%
Stan Dev of Default Pct	6.89%	7.63%	8.30%
Prob(Default Pct > %50)	0.000%	0.011%	0.054%

Parameter Estimation

Choosing the correlation ρ and the correlated diversity score D_ρ

Assume that D is the usual two-moment diversity score for independent, identically distributed representative assets $z_1, .., z_D$. By definition, D is chosen so that

$$Var\left(\frac{1}{D}\sum_{j=1}^{D} z_j\right) = \frac{p(1-p)}{D}$$

is equal to the variance of the actual assets[11]. For clarity, in the remainder of this paper, D will be referred to as the independent diversity score.

Let D_ρ be the correlated diversity score, i.e. the number of Correlated Binomial representative assets x_j (j=1,.., D_ρ) with common pair-wise default correlation ρ. Use the moment matching method to choose ρ and D_ρ. Specifically, for any given value of ρ, choose the correlated diversity score D_ρ so that the variance of the Correlated Binomial representative assets is equal to the variance of the actual assets. The variance of the percentage of defaulted representative assets in the Correlated Binomial is

$$Var\left(\frac{1}{D_\rho}\sum_{j=1}^{D_\rho} x_j\right) = \frac{p(1-p)\left(1+\rho(D_\rho-1)\right)}{D_\rho}$$

10 The examples used in Tables 1 and 2 do not use default stress factors in any of the correlation cases. In practice, Moody's does use default stresses with the BET (independent binomial) but not when correlation is explicitly modeled as in the Correlated Binomial or in simulations.
11 See Appendix II of "Moody's Approach to Rating Multisector CDOs", September 2000.

The single correlation can be estimated as follows[12]. Assume m actual assets each of size N_i with default probability P_i and survival probability $Q_i = 1 - P_i$. Each pair has default correlation ρ_{ij}. Then estimate the single correlation ρ by the default weighted average correlation

$$\overline{\rho} = \frac{\sum_{i=1}^{m}\sum_{j=i+1}^{m}\rho_{ij}N_iN_jP_iP_j}{\sum_{i=1}^{m}\sum_{j=i+1}^{m}N_iN_jP_iP_j}$$

and choose D_ρ, the correlated diversity score, to match the second moment of the default distribution.

$$D_\rho = \frac{\left(\sum_{i=1}^{m}P_iN_i\right)\left(\sum_{i=1}^{m}Q_iN_i\right)\left(1-\overline{\rho}\right)}{\sum_{i=1}^{m}\sum_{j=1}^{m}\rho_{ij}\sqrt{P_iQ_iP_jQ_j}N_iN_j - \overline{\rho}\left(\sum_{i=1}^{m}P_iN_i\right)\left(\sum_{i=1}^{m}Q_iN_i\right)}$$

If D is already known, then for any given value of ρ, choose D_ρ so that

$$Var\left(\frac{1}{D_\rho}\sum_{j=1}^{D_\rho}x_j\right) = Var\left(\frac{1}{D}\sum_{j=1}^{D}z_j\right)$$

This implies that

$$\frac{p(1-p)\left(1+\rho(D_\rho-1)\right)}{D_\rho} = \frac{p(1-p)}{D}$$

and is equivalent to Expression (4) below[13]

Expression (4): $D_\rho = \dfrac{(1-\rho)D}{1-\rho D}$

This equation relates three quantities D_ρ, D and ρ. If the independent diversity score D is known, the correlated diversity score D_ρ can be calculated by substituting $\overline{\rho}$ for ρ into Expression (4)[14]. Consider the following example of estimating D_ρ and ρ when D is known. Twenty-five actual assets have an average default probability of 5%. Using a two-moment calculator, the independent diversity score[15] of these twenty-five assets has been calculated as D = 10. If the estimated correlation is 2.5%, then from Expression (4), D_ρ = 13. If the estimated correlation is 5%, then from Expression (4), D_ρ = 19. *Table 2* illustrates this example. *Table 2* differs from *Table 1* in that the number of assets in each column of *Table 2* varies and is chosen to keep the standard deviation of the default percentage fixed.

12 Other approaches may be used to estimate the correlation parameter in some specific transactions.

13 Note that for $\rho \geq 1/D$, it is not possible to match the second moment for the given value of ρ.

14 In the special case where all m assets are identical with a common constant correlation ρ, from Appendix II of "Moody's Approach to Rating Multi-sector CDOs", the independent diversity score D = m/(1+(m-1)ρ). Substituting m/(1+(m-1)ρ) for D into Expression (4) will give D_ρ = m.

15 Note how this corresponds to the zero correlation column of Table 1. In Table 1 the number of assets is fixed at ten for all three correlation cases.

Table 2

Correlation =	0.00%	2.50%	5.00%
Correlated Diversity Score =	10	13	19
Defaults		**Probabilities**	
0	59.87%	55.96%	50.31%
1	31.51%	28.98%	25.50%
2	7.46%	10.64%	12.67%
3	1.05%	3.27%	6.16%
4	0.10%	0.88%	2.93%
5	0.006%	0.212%	1.358%
6	0.000%	0.045%	0.612%
7	0.000%	0.009%	0.268%
8	0.000%	0.001%	0.113%
9	0.000%	0.000%	0.046%
10	0.000%	0.000%	0.018%
11		0.000%	0.007%
12		0.000%	0.002%
13		0.000%	0.001%
14			0.000%
15			0.000%
16			0.000%
17			0.000%
18			0.000%
19			
Expected Default Pct	5.00%	5.00%	5.00%
Stan Dev of Default Pct	6.89%	6.89%	6.89%
E(Loss at %20 attachment)	0.010%	0.031%	0.050%

The last row of *Table 2* is E(L), the expectation of a simplified loss function L is defined as

$$L = Max\left(0, \frac{X(1-R)/D_\rho - A}{1-A}\right).$$

Here X is the number of defaults, R is the recovery, and A is the lower attachment point. The upper attachment point is assumed to be one. The calculation E(L) in the last row of *Table 2* assumes that R = 30% and A = 20%. Appendix I contains examples that systematically compare the Correlated Binomial, the independent Binomial and simulations based on normal copulas.

Conclusion

The Correlated Binomial combines the advantages of a relatively simple discrete distribution with an explicit correlation assumption. The Correlated Binomial is a closed form distribution, is simple to describe and is intuitively appealing; however, it also has the same higher probability of multiple defaults that characterize correlated portfolios. Because the Correlated Binomial explicitly models correlation, it does not require default probability stresses used in conjunction with the BET. As with the Moody's application of the normal copula approach, no default probability stress is necessary when applying the Correlated Binomial.

The Correlated Binomial represents another analytical tool that Moody's may apply in our CDO rating process. Most pertinent to deals with low diversity scores and highly correlated assets, the Correlated Binomial offers a relatively simple and reliable alternative to more computationally challenging methods.

APPENDIX I

Example Comparison of the BET, the Correlated Binomial and Normal Simulations

Table 4 (see next page) is a comparison of expected losses for (1) Binomial with current default probability stresses, (2) simulations using Normal copulas (no stress) and (3) the Correlated Binomial (no stress). The hypothetical, uniform portfolios, which are described in *Table 3* (see next page), all consist of identical assets with a single constant default correlation between every pair[16]. These hypothetical portfolios are meant to illustrate cases with low Diversity score and high correlation.

The independent diversity score and the correlated diversity score shown in adjacent columns insure that the standard deviation of the default distribution is constant for all three cases[17]. The Binomial stress assumes that the assets and liabilities are ten years long, i.e. a 5% default probability is stressed as a 500 rating factor and the expected losses determine the liability stress based on ten year geometric mean[18].

For each simulation using a normal copula, the default correlation is converted to an equivalent asset correlation[19] and simulated to estimate[20] the expected loss. For both the simulations and the correlated binomial, the default probability of any single asset is given by the second column of *Table 3* headed Default Probability.

The expected loss calculation for each tranche, A and B (as defined by the attachment points in Table 4 below), assumes immediate asset losses. In each random default scenario, the loss function[21] is

Loss = Min{ 1 , Max [0 , (Default Pct*Severity - Lower AP)/(Upper AP - Lower AP)] }

where AP is the attachment point and Default Pct is the percentage of assets defaulting in the given scenario. Recoveries remain constant at 30%. For the two discrete distributions, the expected loss is calculated by as the sum across each possible outcome[22] of the product of the Loss and the probability of that outcome. For the simulations, the expected loss is estimated by averaging the loss results of all the simulated default trials.

In *Table 4* the expected loss from the Correlated Binomial and from the normal copula approach are similar for both tranches across all four portfolios. It is the independent Binomial (i.e. the BET) that differs from the other two approaches. The differences widen as the correlation increases[23] because, as demonstrated by this example, the expected loss for the independent binomial does not change as the correlation increases if the diversity score remains constant. The increasing differences in expected losses are a direct result of the fact that two of the approaches incorporate correlation explicitly into the representative portfolios and one does not.

It is important to note that this example was constructed specifically to illustrate the impact on expected losses of CDO tranches supported by high correlation, low diversity assets. All three distributions, the BET, the Correlated Binomial and the normal copula approach, would have expected liability losses that are much more similar to one another for less correlated, more diverse portfolios.

16 The example does not consider the impact of variation in default probabilities of the actual assets. The correlated binomial, like the BET, is best suited for portfolios of assets of similar credit quality.

17 Consider the two sets of assets representing portfolio (1). One has ten independent assets each with a 5% default probability. That representative portfolio has a percentage default distribution with the same standard deviation, 6.89%, as thirteen assets each with a 5% default probability and 2.5% default correlation between each pair.

18 For example, in portfolio 1, tranche A, the expected loss of 0.019% is just above the geometric mean (0.00173%) for a ten year **Aaa** and **Aa1** in the idealized expected loss table. This means that the stress used is the ratio of the **Aa1** stress (1.45) to an interpolated stress (1.213) between **Baa2** and **Baa3** based on the 500 rating factor of the assets. The ratio of these two stresses increases the default probability in the BET from 5% to 5.98% as reported in the last column of *Table 4*.

19 The asset correlation is chosen so that the probability that any two assets default is the same for the two methods: (1) simulations using normal copulas and (2) the correlated binomial. The default distributions of the two methods differ only for the probability of three or more defaults.

20 Each simulation used 1,000,000 trials. The accuracy resulting expected loss estimates can be gauged by the following sample standard deviations of those expected loss estimates: 1A 0.0005%, 1B 0.0110% , 2A 0.0008%, 2B 0.0123%, 3A 0.0004%, 3B 0.0107%, 4A 0.0006%, 4B 0.0118%.

21 The loss function has been changed slightly from the one used to generate *Table 2* in order to accommodate an upper attachment point.

22 The number of outcomes is the number of representative assets plus one for the case of no defaults.

23 Portfolios 1 and 2 have the same diversity score of ten (independent asset total) but different correlations. As the correlation increases from portfolio 1 to portfolio 2, the expected loss increases for the two explicitly correlated distributions but not for the independent Binomial. The same conclusion holds for portfolios 3 and 4.

Table 3
Portfolios

Portfolio	Default Probability	Default Correlation	Asset Correlation	Independent Asset Total	Correlated Asset Total	St Dev of Default Distribution
1	5.0%	2.5%	9.50%	10	13	6.89%
2	5.0%	5.0%	17.75%	10	19	6.89%
3	10.0%	6.25%	16.00%	8	15	10.61%
4	10.0%	10.0%	24.25%	8	36	10.61%

Table 4
Expected Losses

Portfolio	Tranche	Exp Loss: Independent Binomial	Exp Loss: Correlated Binomial	Exp Loss: Normal Simulation	Lower Attachment	Upper Attachment	Independent Binomial Stressed Default Prob
1	A	0.019%	0.0307%	0.0317%	21.00%	100.00%	5.98%
1	B	1.275%	1.7773%	1.7417%	15.00%	21.00%	5.19%
2	A	0.019%	0.0503%	0.0588%	21.00%	100.00%	5.98%
2	B	1.275%	1.8412%	1.8786%	15.00%	21.00%	5.19%
3	A	0.009%	0.0112%	0.0133%	40.50%	100.00%	12.76%
3	B	1.868%	2.5445%	2.5757%	17.50%	40.50%	10.46%
4	A	0.009%	0.0214%	0.0227%	40.50%	100.00%	12.76%
4	B	1.868%	2.7139%	2.7923%	17.50%	40.50%	10.46%

APPENDIX II

Derivation of the Correlated Binomial default distribution

In the Correlated Binomial, the probability of no defaults and n survivals is

$$E\left[\prod_{j=1}^{n}(1-x_j)\right] \quad = \quad 1+\sum_{j=1}^{n}(-1)^j C(n,j)\prod_{i=1}^{j} p_i$$

As an illustration, consider the special case of n=3.

$$E\left[\prod_{j=1}^{3}(1-x_j)\right] \quad = \quad E[(1-x_1)(1-x_2)(1-x_3)] \quad = \quad E[1-x_1-x_2-x_3+x_1x_2+x_1x_3+x_2x_3-x_1x_2x_3]$$

$$E\left[\prod_{j=1}^{3}(1-x_j)\right] \quad = \quad 1-E(x_1)-E(x_2)-E(x_3)+E(x_1x_2)+E(x_1x_3)+E(x_2x_3)-E(x_1x_2x_3)$$

Because all the assets are identically distributed,

$E(x_1)=E(x_2)=E(x_3) \quad and \quad E(x_1x_2)=E(x_1x_3)=E(x_2x_3)$.

Therefore

$$E\left[\prod_{j=1}^{3}(1-x_j)\right] \quad = \quad 1-3E(x_1)+3E(x_1x_2)-E(x_1x_2x_3)$$

For any value of n, the probability of no defaults and n survivals is

$$E\left[\prod_{j=1}^{n}(1-x_j)\right] \quad = \quad E[(1-x_1)\ \ldots\ (1-x_n)]$$

Using a standard polynomial expansion

$$E\left[\prod_{j=1}^{n}(1-x_j)\right] \quad = \quad 1+\sum_{j=1}^{n}(-1)^j C(n,j)E\left(\prod_{i=1}^{j} x_i\right)$$

Recall that from the definition of p_j

$$p_j = E(x_j \mid x_1=1, x_2=1, \ldots, x_{j-1}=1) \quad for\ j=1,..,n$$

that

$$\prod_{j=1}^{k} p_j = E\left(\prod_{j=1}^{k} x_j\right) \quad for\ k=1,..,n$$

which implies by substitution that

$$E\left[\prod_{j=1}^{n}(1-x_j)\right] = 1+\sum_{j=1}^{n}(-1)^j C(n,j)\prod_{i=1}^{j} p_i$$

For k > 0, the probability that assets 1 to k default and assets k+1 to n survive is

$$E\left[\prod_{j=1}^{k} x_j \prod_{j=k+1}^{n}(1-x_j)\right] = E\left[x_1 \ldots x_k (1-x_{k+1}) \ldots (1-x_n)\right]$$

$$E\left[\prod_{j=1}^{k} x_j \prod_{j=k+1}^{n}(1-x_j)\right] = \sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j) E\left(\prod_{i=1}^{j+k} x_i\right)\right] = \sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\prod_{i=1}^{j+k} p_i\right]$$

This scenario, assets 1 through k default and assets k+1 to n survive, is only one specific ordering of k defaults and n-k survivals. Another specific ordering is that assets one through n-k survive and assets n-k+1 to n default. There are C(n, k) such specific orderings. Each is equally probable because the assets are identical. Therefore, the probability of k defaults and n-k survivals in any order is

$$C(n,k) E\left[\prod_{j=1}^{k} x_j \prod_{j=k+1}^{n}(1-x_j)\right] = C(n,k)\sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\prod_{i=1}^{j+k} p_i\right].$$

Relationship to BET

It was previously stated that the BET is the special case of the Correlated Binomial when $\rho = 0$. From the definition of p_j , if $\rho = 0$, then $p_j = p$ for j=1,..,n. If $\rho = 0$, the following steps show that the probability of k defaults and n-k survivals in any order is the independent binomial probability distribution as used in the BET.

$$C(n,k)\sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\prod_{i=1}^{j+k} p_i\right] = C(n,k)\sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\ p^{j+k}\right].$$

$$= C(n,k) p^k \sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)\ p^j\right] = C(n,k) p^k (1-p)^{n-k}.$$

APPENDIX III

Computational Issues

This paper began with three assumptions that define the Correlated Binomial, which as was mentioned in a prior footnote, is more properly a correlated Binomial. Assumptions 1 and 2 would be common to any correlated Binomial. It is Assumption 3 that is specific to Moody's Correlated Binomial. It states that conditional correlations are constant as k increases where k is the number of known defaults with no survivals. It can be viewed as an assumption about the probability that all the assets in a given portfolio will default, i.e. an assumption on the probability of k defaults of k assets. Notice that as stated previously in the section on the definition of the Correlated Binomial,

$$\prod_{j=1}^{k} p_j = E\left(\prod_{j=1}^{k} x_j\right) \quad for\ k = 1,..,n$$

The product $\prod_{j=1}^{k} p_j$ is the probability of k defaults in a portfolio of k assets. A different correlated Binomial will result if a different assumption is made in place of Assumption 3. As an example, the normal copula could be used to specify $\prod_{j=1}^{k} p_j$, the probability of k defaults in k assets. Using the normal copula, the formula for the probability of k defaults out of n uniform, correlated assets is

$$C(n,k)\sum_{j=0}^{n-k}\left[(-1)^j C(n-k,j)E\left(\prod_{i=1}^{j+k} x_i\right)\right],$$

where the product $E\left(\prod_{j=1}^{j+k} x_j\right)$ is the probability that j+k correlated standard normal random variables fall below the appropriate threshold level (determined from the default probability). This distribution would be very similar to Moody's Correlated Binomial as was indicated by the expected loss results[24] in *Table 4*. Other copulas could be used in the same way.

However, if calculations are performed in standard double precision floating point arithmetic, as in most spreadsheet software, the closed form formula above will result in accumulating computational errors as the number of assets and/or the correlation increases. This will be true for Moody's Correlated Binomial, normal copulas, or other copulas because this closed form formula is a succession of differences between products of increasingly small numbers multiplied by increasingly large numbers.

One advantage of Moody's Correlated Binomial compared to other possible correlated Binomials is that the simplicity of Assumption 3 allows for a relatively simple algorithm to produce enhanced arithmetic precision and increased computational tractability. The algorithm below computes Moody's Correlated Binomial probabilities using the following simple scheme that breaks the calculation down to a series of differences and products involving numbers between zero and one[25].

First, recall the definition

$$p_j = E(x_j \mid x_1 = 1, x_2 = 1,..., x_{j-1} = 1) \quad for\ j = 1,..,n$$

24 The normal copula was simulated to produce the results in *Table 4* but the same probabilities could have been calculated by numerically integrating $E(\prod x_j)$ and using the result in the formula above.

25 This algorithm is mathematically equivalent to the closed form distribution formula.

Also recall that $p_1 = p$ and that $p_2 = p + (1-p)\rho$ which implies that the probability that assets 1 and 2 default, $P(x_1=1 \cap x_2=1)$, is known.

Similarly, the assumption that

$$p_{j+1} = p_j + (1-p_j)\rho \quad for\ j = 1,..,n-1$$

implies that the probability that assets 1 to n default, $P(x_1=1 \cap \ldots \cap x_n=1)$, is also known and can be computed as a series of products of probabilities between zero and one:

$P(x_1=1 \cap \ldots \cap x_n=1) = \prod_{j=1}^{k} p_j$.

Special Case of Two Assets

Since $P(x_1=1 \cap x_2=1) + P(x_1=1 \cap x_2=0) = P(x_1=1)$,
the probability of one default and one survival is
$P(x_1=1 \cap x_2=0) = P(x_1=1) - P(x_1=1 \cap x_2=1)$.

Similarly, the probability of two survivals of two assets is
$P(x_1=0 \cap x_2=0) = P(x_1=0) - P(x_1=0 \cap x_2=1)$

Since the default distributions of the assets are identical, the probability of one default and one survival is the same regardless of order, i.e. $P(x_1=1 \cap x_2=0) = P(x_1=0 \cap x_2=1)$ and therefore
$P(x_1=0 \cap x_2=0) = P(x_1=0) - P(x_1=1 \cap x_2=0)$

Let $P_{j,n}$ represent the probability of j defaults of n assets in one specific order. For example, $P_{1,2} = P(x_1=1 \cap x_2=0) = P(x_1=0 \cap x_2=1)$. Then, as was just shown
$P_{1,2} = P_{1,1} - P_{2,2}$ and $P_{0,2} = P_{0,1} - P_{1,2}$

General Case of n Assets

The computational algorithm relies on the two following relationships.

$P_{k,k} = \prod_{j=1}^{k} p_j$

$P_{j-1,k} = P_{j-1,k-1} - P_{j,k}$

$P_{1,1}$ and p_1 are known. Any subsequent values of $P_{j,k}$ can be found from the following[26].

$P_{0,1} = 1 - P_{1,1}$
For k = 2 to n
  $p_k = p_{k-1}(1-\rho) + \rho$
  $P_{k,k} = p_k P_{k-1,k-1}$
  For j = k-1 to 0 step -1

26 Note that the only repeated arithmetic operations are products and differences of probabilities.

```
    P(j, k) = P(j, k-1) - P(j+1, k)
  Next j
Next k
```

It is important to note that the output of this algorithm is the probability of j defaults of n assets in a specific order. To get the probability of j defaults of n assets in any order, multiply by the combination of n objects selected j at a time, C(j, n) where $C(j,n) = \dfrac{j!}{n!(j-n)!}$.

The full algorithm is given below in Visual Basic[27]. It consists of a top level routine *IteratePrecisionProb()* that uses the inductive algorithm outlined above. *IteratePrecisionProb()* repeatedly calls *Difference,* a routine that performs enhanced precision differences, and *Product,* a routine that performs enhanced precision multiplication. Both depend on the fact that the numbers involved are all between zero and one. In addition there are two routines, *XLongToDouble* and *DoubleToXLong*, to translate numbers from double precision floating point to enhanced precision and back. In the code below, enhanced precision probabilities are stored in variables of the data type XLong, an array of long integers.

Algorithm for Enhanced Precision Calculation of the Correlated Binomial

Comments are enclosed in quotation marks "".

Declaration of Global Variables and Constants

```
Const ExponentSize = 8
Const Size = 10 ^ 8
Const HalfSize = 10 ^ 4
Const MaxDim = 100
Const MaxLength = 8
Const Precision = 10    "Number of Longs used to represent a probability"

Type XLong
  Part(Precision) As Long
End Type
```

"XLong data type is used to represent a number between 0.0 and 1.0 as an array of Longs. Each Long is used to hold eight base 10 digits."

```
Dim IntCorr As Long
```

Public Sub IteratePrecisionProb()

" This is the top level routine that calculates correlated Binomial probabilities by taking successive differences."

27 Although written here in VB, this is only a general algorithm. Moody's assumes no responsibility for any implementation issues that may arise for specific systems.

```
Dim DefProb As Double
Dim Correlation As Double
Dim Association As Integer
Dim Diversity As Integer
Dim i, j As Integer
Dim Cond(MaxDim) As Double
Dim XCond(MaxDim) As XLong
Dim Prob(MaxDim) As Double
Dim OneMinusCorr As XLong

“XProb(m,n) holds the probability of m defaults on n trials. It is calculated iteratively.”

Dim XProb(MaxDim, MaxDim) As XLong

Range("ErrorMsg") = ""
Range("OutputArea").Clear

DefProb = Range("DefaultProbability")
If DefProb = 0 Then DefProb = 1/size
If DefProb = 1# Then DefProb = 1# - 1/size
If DefProb < 0# Then
Range("ErrorMsg") = "Default Probability cannot be below zero."
Exit Sub
End If
If DefProb > 1# Then
Range("ErrorMsg") = "Default Probability cannot be greater than one."
Exit Sub
End If

Correlation = Range("Correlation")
If Correlation = 0 Then Correlation = 1/size
If Correlation = 1# Then Correlation = 1# - 1/size
If Correlation < 0# Then
Range("ErrorMsg") = "Correlation cannot be below zero."
Exit Sub
End If
If Correlation > 1# Then
Range("ErrorMsg") = "Correlation cannot be greater than one."
Exit Sub
End If
```

```
Diversity = Range("NumAssets")
IntCorr = CLng(Correlation * Size)
Calculate

Prob(1) = DefProb
XCond(1) = DoubleToXLong(DefProb)
OneMinusCorr = DoubleToXLong(1 - Correlation)
XProb(1, 1) = DoubleToXLong(DefProb)
XProb(0, 1) = DoubleToXLong(1 - DefProb)

For i = 2 To Diversity
"The next two statements calculate the conditional prob of i-th default given i-1 defaults
The distribution is assumed to have a conditional correlation between asset i and asset i-1 that is constant given i-2 defaults."

   Call Product(XCond(i - 1), OneMinusCorr, XCond(i))
   XCond(i).Part(1) = IntCorr + XCond(i).Part(1)
   Call Product(XCond(i), XProb(i - 1, i - 1), XProb(i, i))
   Prob(i) = XLongToDouble(XProb(i, i))
Next i

For i = 2 To Diversity
For j = i - 1 To 0 Step -1
   Call Difference(XProb(j, i - 1), XProb(j + 1, i), XProb(j, i))
Next j
Next i

For i = 0 To Diversity
   Range("ProbOut").Offset(0, i) = XLongToDouble(XProb(i, Diversity))
Next i
Calculate
End Sub
```

Public Sub Difference(Big As XLong, Small As XLong, Result As XLong)

```
"Calculates the Difference using Arbitrary Precision"

Dim i As Integer

Erase Result.Part
```

```
Result.Part(1) = Big.Part(1) - Small.Part(1)
For i = Precision to 2 step -1
  Result.Part(i) = Big.Part(i) - Small.Part(i) + Result.Part(i)
  If Result.Part(i) < 0 Then
    Result.Part(i - 1) = Result.Part(i - 1) - 1
    Result.Part(i) = Size + Result.Part(i)
  End If
Next i
End Sub
```

Public Sub Product(Factor1 As XLong, Factor2 As XLong, Result As XLong)

```
"Calculates the Product using Arbitrary Precision."

Dim i, j, IntSum As Integer
Dim LoResult, HiResult, x As Long

Erase Result.Part
For IntSum = Precision To 2 Step -1
For j = 1 To IntSum - 1
i = IntSum – j
Call PartMultiplier(Factor1.Part(i), Factor2.Part(j), LoResult, HiResult)
Result.Part(i + j) = LoResult + Result.Part(i + j)
If Size < Result.Part(i + j) Then
    x = Int(Result.Part(i + j) / Size)
    Result.Part(i + j) = Result.Part(i + j) - x * Size
    Result.Part(i + j - 1) = Result.Part(i + j - 1) + x
End If
Result.Part(i + j - 1) = HiResult + Result.Part(i + j - 1)
If Size < Result.Part(i + j - 1) Then
    x = Int(Result.Part(i + j - 1) / Size)
    Result.Part(i + j - 1) = Result.Part(i + j - 1) - x * Size
    Result.Part(i + j - 2) = Result.Part(i + j - 2) + x
End If
Next j
Next IntSum

End Sub
```

Public Sub PartMultiplier(Factor1 As Long, Factor2 As Long, LowerResult As Long, UpperResult As Long)

```
Dim Fac1(2), Fac2(2), Result(2) As Long
Dim Middle, UpperMiddle, LowerMiddle As Long
Dim x As Long

Fac1(1) = Int(Factor1 / HalfSize)
Fac1(2) = Factor1 - Fac1(1) * HalfSize
Fac2(1) = Int(Factor2 / HalfSize)
Fac2(2) = Factor2 - Fac2(1) * HalfSize
Result(1) = Fac1(1) * Fac2(1)
Result(2) = Fac1(2) * Fac2(2)
Middle = Fac1(1) * Fac2(2) + Fac1(2) * Fac2(1)
UpperMiddle = Int(Middle / HalfSize)
LowerMiddle = Middle - UpperMiddle * HalfSize
Result(2) = Result(2) + LowerMiddle * HalfSize
x = Int(Result(2) / Size)  ' if there is (8 digit) overflow
Result(2) = Result(2) - x * Size
Result(1) = Result(1) + x
Result(1) = Result(1) + UpperMiddle
UpperResult = Result(1)
LowerResult = Result(2)
End Sub
```

Public Function DoubleToXLong(InDouble As Double) As XLong

```
Dim x, y, e As Double
Dim q as XLong
Dim Ylong As Long
Dim i,j,n, Modulus, DivResult, PartIndex, DigitPlace As Integer
Dim Digits(MaxLength) As Integer

x = InDouble
e = Log(x) / Log(10)
n = -Int(e)
y = x * 10 ^ (n + MaxLength)
YLong = Int(y)

For j = 1 To MaxLength
Digits(j) = Int(YLong * 10 ^ (j - MaxLength - 1))
```

```
YLong = YLong - Digits(j) * 10 ^ (MaxLength - j + 1)
Next j

DivResult = (n - 1) \ ExponentSize
Modulus = n Mod ExponentSize
If Modulus = 0 Then Modulus = ExponentSize
PartIndex = DivResult + 1
DigitPlace = Modulus

For j = 1 To MaxLength
q.Part(PartIndex) = q.Part(PartIndex) + Digits(j) * 10 ^ (ExponentSize - DigitPlace)
If DigitPlace = ExponentSize Then
    DigitPlace = 1
    PartIndex = PartIndex + 1
Else
    DigitPlace = DigitPlace + 1
End If
Next j

DoubleToXLong = q
End Function
```

Public Function XLongToDouble(InXLong As XLong) As Double

```
Dim j As Integer
Dim q As XLong
Dim x, CurrentSize As Double

q = InXLong
CurrentSize = Size
For j = 1 To Precision
    x = x + q.Part(j) / CurrentSize
    CurrentSize = CurrentSize * Size
Next j

XLongToDouble = x

End Function
```

Doc ID# SF42036