Three Essays on the Risk and Distribution of a Portfolio's Future Losses

by

Wei He

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Rotman School of Management
University of Toronto

© Copyright by Wei He (2001)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-63706-9
Abstract

Three Essays on the Risk and Distribution of a Portfolio’s Future Losses
Doctor of Philosophy
Graduate Department of Rotman School of Management
University of Toronto

This Ph.D. dissertation contains three individual and internally related essays.

The first essay applies the least-squares Monte-Carlo (LSM) methodology to derive the distribution of the exotic option values at a future time. LSM presents a powerful statistical procedure that efficiently yields derivative distributions for exotic options that do not possess analytic solutions. By means of several examples, using options with closed-from solutions, this essay demonstrates the ability of LSM to produce excellent estimates of derivative distribution at a reasonable computational cost.

The second and third essays compare two of the major credit risk portfolio models used by two prominent financial companies: J.P. Morgan’s CreditMetrics and Credit Swiss First Boston’s CreditRisk+. The second essay compares the two models from a methodological and an empirical point of view. Factor Analysis is utilized to link the different input data employed by these two models. The third essay creates a hypothetical world in which the true transition matrices are known so that a benchmark distribution of portfolio loss is derived to evaluate the model’s performance. The results suggest that despite the fact that the recommendations made by each approach to a financial institution trying to determine how much economic capital to hold is different, these two models perform equally well when credit-rating-change risk is eliminated from the CreditMetrics approach.
I am most grateful to my supervisor Alan White for his valuable advice, suggestions, and encouragement. His supervision and guidance were invaluable in all aspects of this dissertation. I thank Bram Cadsby for his insightful comments and discussions. I also thank my thesis committee Laurence Booth, Peter Carayannopoulos, Paul Halpern and John Hull for their helpful comments.
Introduction

For risk management purposes, financial institutions are interested in calculating a complete probability distribution for losses that could be experienced in the real world at a future time.

Predicting future events with certainty is not possible. In a financial market, uncertainty that is associated with the value of a position is risk. There are primarily four different sources of risk: market risk, credit risk, operational risk and liquidity risk. Market risk is the uncertainty resulting from changes in market conditions, credit risk results from changes in the ability of a counterparty to meet its obligations, while operational risk is the name given to uncertainty resulting from a breakdown in communications, information or transactional processing, or legal/compliance issues. Liquidity risk is defined as uncertainty resulting from the inability of a firm to fund its illiquid assets. In order to achieve effective measurement and management of each of these types of risk, Risk Management Analysis aims to apply the methods of empirical science to estimate possible future events according to the laws of probability and statistics.

Since future losses cannot be known with absolute certainty, Risk Management Analysis gauges risk by focusing on estimates of the distributions from which future events may be drawn. Knowledge of this distribution allows the analyst to determine the probability of suffering a loss which exceeds a given threshold, thus in turn allowing prediction of the largest future loss at a given confidence level.

Bank capital is a bank's cushion against possible losses emanating from its other assets. Capital must be held in the form of liquid and secure assets. When losses arising
from its other assets exceed the total amount of capital, the bank defaults and the excess loss is passed along to the creditors. Therefore, it is important for the bank to hold enough capital such that the probability of default is minimized.

By combining capital reservation requirements with the distributions of possible losses and analysis of market behavior over a long historical period, financial institutions calculate their own estimates of capital levels required to guard against risk. Value at Risk (VaR) measures consider two basic parameters, the confidence level and the time horizon, which define the relevant regions of the distributions of possible loss. For market risk, for example, the required capital is calculated by multiplying the 10-day 99% VaR by three. The 10-day 99% VaR of a portfolio is that loss over 10 days which there is a 1% chance of exceeding. With respect to credit risk, the economic capital is determined by the unexpected loss of the portfolio due to credit risk. The unexpected loss is computed by subtracting the expected loss from the 99th percentile of the loss distribution over a one-year time horizon.

This Ph.D. dissertation, focusing on risk management in the market risk and credit risk area, contains the following three individual and internally related essays:

1) An Efficient Numerical Method of Determining the Distribution of an Exotic Option’s Future Value Using Least-Squares Monte-Carlo Simulation Approach;

2) Comparison of Two Credit Risk Portfolio Models: CreditMetrics and CreditRisk+;

and

3) CreditMetrics versus CreditRisk+: A Comparison of Two Credit Risk Portfolio Models When Probabilities of Rating Migrations Are Known.
The first essay sets foot on the market risk field specifically addressing the issue of obtaining the distribution of the future loss for exotic options of which there are no analytic solutions to pricing. A great deal of attention has been devoted to the development of efficient numerical procedures. This essay applies the least-squares Monte-Carlo (LSM) methodology to derive the distribution of the exotic option values at a future time. By combining the least-squares regression technique with the pure Monte-Carlo simulation method, LSM approach requires only least-squares regressions and circumvents the computational difficulty of the Monte-Carlo procedure when deriving the expected future derivative value given certain future underlying values.

The second and third essay, in contrast, investigate the methodologies of deriving the distribution of a portfolio’s future loss when credit risk is the primary driving force of the uncertain future position value. These two essays compare two of the major credit risk portfolio models used by two prominent financial companies: J.P. Morgan’s CreditMetrics and Credit Swiss First Boston’s CreditRisk+. The second essay compares the two models from a methodological and an empirical point of view. To facilitate the methodological comparison, I developed a framework to link the different input data employed by these two models. In particular, I showed that if asset values of different counterparties are related to some economic factors, factor analysis may be used to transform the asset correlation data implied by CreditMetrics into the standard deviation data employed by CreditRisk+. The biggest problem when comparing these two models is that we do not know the true transition matrix, nor do we know how it may change from year to year. The data provided by J.P. Morgan are estimates of the true values based on historical data. They are used because the true values are unknown. When such
data are used to compare the two credit risk portfolio models, all that we can learn is how each model would have performed over the historical period on which the estimates are based. We cannot necessarily conclude that the better performing model will continue its superior performance in the future. This problem is compounded by the fact that the historical estimates yield just one estimated transition matrix rather than allowing for changes in the matrix that might occur owing to changing economic conditions over the business cycle. To resolve the problem, I create a hypothetical world in the third essay in which the true transition matrices are known so that a benchmark distribution of portfolio loss required for the comparison can be derived for the purpose of evaluating the model's performance. When implementing the CreditMetrics approach, I used the default correlation data directly to incorporate the joint default behavior. Given the known transition matrices and the implied counterparty migration probabilities, the standard deviation of default rates and the default correlation data can be derived directly from the same set of historical data. This has two advantages. First of all, since credit-rating-change-risk is omitted from the CreditMetrics approach, the default correlation data are sufficient for modeling the joint default behavior. Secondly, CreditMetrics and CreditRisk+ are easily comparable because the data used to represent the joint default behavior of the two models are consistent. This consistency eliminates the risk of modeling errors when transforming the asset correlation data into the standard deviation of default rates for the purpose of comparison. It also eliminates the necessity of finding an appropriate proxy to derive the asset correlations.

These three essays are self-contained, each of which consists of four general sections: Introduction, Methodology/Model, Testing/Results and Conclusion.
An Efficient Numerical Method for Determining the Distribution of an Exotic Option’s Future Value Using Least-Squares Monte-Carlo Simulation Approach
I. Introduction

For risk management purposes, financial institutions are interested in calculating the complete probability distribution for losses that could be experienced at some time in the future.

Predicting future events with certainty is not possible. In a financial market, the uncertainty that is associated with the value of a position is referred to as the risk of the position. There are four primary sources of risk: market risk, credit risk, operational risk and liquidity risk. Market risk is the uncertainty resulting from changes in market conditions; credit risk results from changes in the ability of a counterparty to meet its obligations; operational risk is the name given to uncertainty resulting from a breakdown in communications, information or transactional processing, or legal/compliance issues; and liquidity risk is defined as uncertainty resulting from the inability of a firm to fund its illiquid assets. In order to achieve effective measurement and management of each of these types of risk, Risk Management Analysis aims to apply the methods of empirical science to estimate possible future events according to the laws of probability and statistics.

Since future losses cannot be known with absolute certainty, Risk Management Analysis gauges risk by focusing on estimates of the distributions from which future events may be drawn. Knowledge of this distribution allows the analyst to determine the probability of suffering a loss that exceeds a given threshold. This knowledge of the probability distribution also allows prediction of the largest future loss at any given confidence level.
Bank capital is a bank's cushion against possible losses emanating from its other assets. Capital must be held in the form of liquid and secure assets. When losses arising from its other assets exceed the total amount of capital, the bank defaults and the excess loss is passed along to the creditors. Therefore, it is important for the bank to hold enough capital such that the probability of default is minimized.

By combining capital reservation requirements with the distributions of possible losses and the analysis of market behavior over a long historical period, financial institutions calculate their own estimates of capital levels required to guard against risk. The current practice amongst financial institutions is to base their capital on their Value at Risk (VaR). The behavior of market variables is monitored over a long historical period to determine their statistical properties. These statistical properties are then merged with the institution's portfolio to determine the probability distribution of possible future gains and losses. The capital set aside is then based on some measure of the largest probable loss.

Value at Risk (VaR) has two basic parameters: the confidence level for the largest probable loss, and the time horizon over which forecasts are made. Once these two parameters are chosen, the relevant regions of the distributions of possible loss can then be determined. Under current regulations, for market risk the time horizon is 10 days and the confidence level is 99%. The required capital is calculated by multiplying the 10-day, 99% VaR by three. The 10-day 99% VaR of a portfolio is that loss over a period of 10 days which will only be exceeded with a probability of 1%. For credit risk the time horizon is 1 year and the confidence level is 99%. The economic capital is based on the unexpected loss of the portfolio due to credit risk. The unexpected loss is calculated by
subtracting the expected loss from the 99\textsuperscript{th} percentile of the loss distribution over a one-year time horizon.

The calculation of VaR is conceptually straightforward and in practise it is usually a fairly simple calculation. The value of most of the financial institution’s portfolio is a linear function of the market variables. For these assets and liabilities the translation from the probability distribution of the market variables to the probability distribution of the portfolio is very simple. For assets and liabilities whose value is not a linear function of the market variables, usually derivatives, the translation from the distribution of the market variables to the distribution of values may be more complicated.

For a derivative of which an analytic pricing formula is available, obtaining the distribution of the derivative’s future value is fairly straightforward. The analytic pricing formula is the connection that maps the distribution of the future underlying variable value to the distribution of the future derivative value. For derivatives for which an analytic pricing formula is not available, a great deal of attention has been devoted to the development of efficient numerical procedures for deriving the derivative distribution. Among the numerical procedures developed to calculate such a derivative value, the most common methodologies are the tree or lattice approach\textsuperscript{*}, and the Monte-Carlo simulation approach.

These methods often slow down the calculation of VaR since in order to determine the probability distribution of the derivative value it may be necessary to recompute the derivative value for many different values of the market variables. In some cases the tree methodology allows us to read the distribution of values directly from the tree. As long as the underlying value of the market variable is constructed forward along
the tree at each node and the corresponding derivative value is derived backward along the tree, the distribution of the future derivative value can be abstracted directly off the tree. In order to develop much accuracy in the forecast distribution of values, a very dense tree must be used. A great number of exotic options cannot be priced through the tree method and consequently, the distribution of these derivatives' future value cannot be derived in this way.

Path dependent and other similar derivatives must be valued by Monte Carlo simulation. In this case the future behavior of the underlying market values is simulated, and given the future values of the market variables, the payoff on the derivative is computed. This provides an unbiased estimate of the value of the derivative. If this process is repeated many times we get many such unbiased estimates of the value. If the simulations are independent, the errors in the estimates should also be independent. By averaging the results we can get a fairly accurate estimate of the value of the derivative.

Monte Carlo simulation is a powerful and robust way of estimating the value of a derivative. However, since many simulations are required it is quite slow. As a result, it is usually computationally impractical to derive the distribution of a derivative's future value because it would require a large number of Monte-Carlo simulations, each one associated an alternative initial value of the market variables. This paper investigates a methodology to overcome this computational obstacle by combining a regression with the Monte-Carlo simulation that is used to compute the current value of the derivative.

The intuition behind this method is simple. In a Monte Carlo simulation the path followed by the underlying market variable from today to the derivative maturity is simulated. As the variable is simulated toward the maturity date, the effect that it has on

---

*Finite difference approaches are equivalent to the tree approach.*
the value of the derivative is calculated until at the maturity date the total value of the derivative (contingent on the market variables following that particular path) is known. This value is then discounted back along the simulated path to the present producing one estimate of the value of the derivative. The computed value is the true value of the derivative plus an error term. This error term is the difference between the true derivative value and the ex-post realized derivative value for this specific simulation path. By repeating such simulation procedure many times, a series of error terms is then generated. These error terms have zero mean and are independent. If the simulation outcomes are averaged across a large number of simulations the error terms should average out close to zero so the overall average is close to the true value of the derivative.

In carrying out this calculation of the current value of the derivative, we are also producing a set of estimates of the value of the derivative at every future time $t$ ($t$ is a point between today and the maturity) for a wide range of values of the underlying market variables. The least-squares-Monte-Carlo (LSM) procedure is a calculation-efficient method that is used to determine the expected value of a derivative at a future time as a function of the underlying market variables. The relationship between the value of the derivative and the value of the market variables is found by regressing the ex-post realized derivative value at that future time on functions of the corresponding underlying variable across all the simulations. LSM effectively produces the conditional distribution of exotic derivatives such as path-dependent derivatives and derivatives with American-style exercise features. The implementation requires no more than simple least squares regressions.

---

* If interest rates are deterministic we can just discount the final value back to the present in a single calculation. When interest rates are stochastic we must discount at each time step using the interest rate that
The rest of the paper proceeds as follows. Section II describes the methodology. Section III demonstrates a detailed example to implement the method. Section IV conducts various tests to verify the methodology, including a European put option on a stock, a European call option on a pure discount bond and a European down-and-out barrier option. Section V concludes the paper.

II. Methodology

A sample of $n$ observations on a variable, denoted $x_1, x_2, \ldots, x_n$, is a random sample if the $n$ observations are drawn independently from the same population, or its probability distribution. Assume the mean of this probability distribution is $\mu_x$ and the variance is $\sigma_x^2$. The Fundamental Theorem of Statistics states that if we sample randomly with replacement from a population, the empirical distribution function is consistent estimate of the population distribution function. Combining this theory with Asymptotic Theory that applies to the case in which the sample size is infinitely large, we come to the following conclusion. When $n$ is sufficiently large, $\mu_x$ and $\sigma_x^2$ can be estimated as follows:

$$\mu_x = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

and

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_x)^2.$$
These expressions are based on the simple notion that when we draw a sample from a population with a sufficiently large sample size, the sample distribution is a good proxy for the population distribution.

Assume there is another variable $y$ that depends on $x$, and the mean and variance of $y$'s probability distribution are $\mu_y$ and $\sigma_y^2$ respectively. Assume that the function of $y$ on $x$ is continuous and single-valued and the functional form is known, say $y = f(x)$. For each $x_i$ in the random sample, there is a corresponding value for $y_i$ that is mapped by $y_i = f(x_i)$. When $n$ is sufficiently large, the sample distribution of $y_i$ is a good proxy for its population distribution. The mean $\mu_y$ and variance $\sigma_y^2$ can be estimated by the following:

$$
\mu_y = \frac{1}{n} \sum_{i=1}^{n} y_i = \frac{1}{n} \sum_{i=1}^{n} f(x_i),
$$

and

$$
\sigma_y^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \mu_y)^2 = \frac{1}{n} \sum_{i=1}^{n} (f(x_i) - \mu_y)^2.
$$

When the functional form of $y$ on $x$ is not known, a regression model can be utilized to determine the relationship between the dependent variable $y$ and the independent variable $x$. Draw a random sample from the dependent variable population $y_1, y_2, \ldots, y_n$, and another random sample from the independent variable population $x_1, x_2, \ldots, x_n$. The regression can be expressed in the following way,

$$
y_i = \beta_1 \phi_1(x_i) + \beta_2 \phi_2(x_i) + \cdots + \beta_m \phi_m(x_i) + u_i \quad u_i \sim N(0, \sigma_u^2) \quad i = 1, 2, \ldots, n, \quad (1)
$$

8
where \( \varphi_k(x_i), \ k = 1, 2, \ldots, m, \) are given functions of \( x_i, \) \( \beta_k, \ k = 1, 2, \ldots, m, \) are coefficients that define the function \( f(x) = \sum_{k=1}^{m} \beta_k \varphi_k(x_i). \) \( N(\cdot, \cdot) \) is the cumulative density function of a normal distribution. Note that, when \( n \) is sufficiently large, there always exists some function \( f(x) \) such that \( u_i \) is asymptotically distributed as \( N(0, \sigma_u^2) \) no matter what distributions \( x_i \) and \( y_i \) follow. The function \( f(x) \) represents the conditional mean of \( y_i \) which is the mean of \( y_i \) conditional on the value \( x_i. \) This is a standard approach to model how \( y_i \) varies with the value of \( x_i. \) Equation (1) can thus be expressed in a different form as

\[
y_i = E[y_i|x_i] + u_i.
\]  

(2)

Representing the estimated regression parameters as \( b_k, \sum_{k=1}^{m} b_k \varphi_k(x_i) \) is then the estimated mean of \( y_i \) conditional on the value \( x_i. \) The regression model determines the specific relationship between the dependent variable and the independent variable. Thus, given a random sample of \( n \) observations on the independent variable \( x_i \) and a random sample of \( n \) observations on the dependent variable \( y_i, \) we are able to obtain a series of conditional mean values each of which is the mean value of the dependent variable, conditional on a specific independent variable. This series of conditional means approximates the dependent variable \( y's \) population distribution.

LSM is a methodology that applies the above concept in a Monte-Carlo simulation framework. Assume the independent variable is an asset price \( S \) and the dependent variable is an option value \( C. \) The value of the option at time \( t \) is a function of \( t \) and the price of the underlying asset price \( S, C(t, S). \) Note that \( S \) can be considered as the
asset price at time \( t \) only for the time being. Shortly I will address a more general scenario in which the choice of \( S \) extends to some function of the path that the underlying asset values follow between time \( 0 \) and \( t \). Given the stochastic process of the asset price and its initial value \( S_0 \), we are able to simulate the distribution of the asset value at a future point in time, \( t, S_i^t, (i = 1, 2, \ldots, n) \). When there is a closed-form solution to the option value, \( C(t, S_i^t) \), the distribution of the option value at the future time \( t \) is given by plugging the simulated asset price into the closed-form formula. When there is no closed-form solution to the option value, the regression model can be utilized to derive the mean of the option value that is conditional on the simulated asset price, \( E[C_i | S_i^t] \). The random sample of the dependent variable in the regression, the option value \( C_i \), is obtained by discounting each option payoff at maturity back to time \( t \) along the original simulation path that goes through \( S_i^t \). A regression is then estimated on this series of option value against functions of the corresponding underlying asset price.

\[
C_i = \beta_1 \varphi_1(S_i^t) + \beta_2 \varphi_2(S_i^t) + \cdots + \beta_m \varphi_m(S_i^t) + u_i, \quad u_i \sim N(0, \sigma^2), \quad i = 1, 2, \ldots, n,
\]

where \( \varphi_k(S_i^t), \quad k = 1, 2, \ldots, m, \) are given functions of \( S_i^t \) and \( \beta_k, \quad k = 1, 2, \ldots, m, \) are coefficients. Defining the estimated option value as \( \hat{C}_i \), then we have

\[
\hat{C}_i = b_1 \varphi_1(S_i^t) + b_2 \varphi_2(S_i^t) + \cdots + b_m \varphi_m(S_i^t) = \sum_{k=1}^{m} b_k \varphi_k(S_i^t)
\]

This series of conditional mean option values estimated through regression approximates the option value distribution at time \( t \).
Consider now a more general scenario in which the derivative value depends on the underlying value not only at time \( t \), but also on its value between time 0 and time \( t \). The path-dependent derivative is a representative example of such a scenario. The value of the path-dependent derivative at time \( t \) is a function of \( t \), the price of the underlying asset at time \( t \), \( S_t \), and some function of the path followed by the asset price between time 0 and \( t \), \( g(t, S_t) \). Different properties associated with different path-dependent derivatives determine \( g(t, S_t) \). It can be the summation of the underlying asset value from time 0 to \( t \); it can be the arithmetic average of the underlying asset value from time 0 to \( t \); or it can be a certain number of maximum or minimum underlying asset values between time 0 and \( t \).

Denote the path-dependent derivative as \( C(t, S_t, g(t, S_t)) \). To derive the distribution of the derivative value at a future time \( t \), firstly, simulate \( n \) paths of the underlying asset value over the life of the derivative. Secondly, for the path \( i \) we obtain the underlying asset value \( S_t^i \) and the function of the path followed by the asset price between time 0 and \( t \), \( g(t, S_t^i) \). Thirdly, obtain the derivative value in the path \( i \) at time \( t \) by discounting the payoff of the derivative at the maturity back to time \( t \) along the path \( i \), \( C(t, S_t^i, g(t, S_t^i)) \).

To simplify the notation, denote it as \( C_i \). Finally, run a regression of the ex-post realized derivative value \( C_i \) on a function of \( S_t^i \) and \( g(t, S_t) \) as follows,

\[
C_i = F(S_t^i, g(t, S); \beta) + u_i, \quad u_i \sim N(0, \sigma_u^2), \quad i = 1, 2, \ldots, n,
\]

where \( F(S_t^i, g(t, S); \beta) \) is the regression function of the variables \( S_t^i \), the function \( g(t, S_t^i) \) and the unknown parameters \( \beta \). The distribution of the derivative value at time \( t \) is given by
\[ \hat{C}_i = F\left(S_i^t, g\left(t, S_i^t\right); \beta\right), \quad i = 1, 2, \ldots, n. \]

**Functional Form**

So far I have not addressed the issue regarding the functional form of the regression. The regression function describes how the dependent variable varies with the independent variable. Since there is no analytic formula between the dependent variable and the independent variable, polynomials are utilized to approximate the functional form. The Taylor polynomial is one of the fundamental building blocks of this type of numerical analysis.

The Taylor series of a given function \( f(x) \) about a chosen point \( x_0 \) is defined as the infinite series

\[
f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \cdots
\]

\[
= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n.
\]

The purpose of Taylor series is to represent the given function \( f(x) \). Whether the Taylor series converge to \( f(x) \) on some \( x \) interval depends on the following three conditions:

1. The function \( f(x) \) is infinitely differentiable at \( x_0 \) so that the coefficients \( \frac{f^{(n)}(x_0)}{n!} \) in the Taylor series exist;

2. The Taylor series converges in some interval \( |x-x_0| < R \), for \( R > 0 \);

3. The sum of the Taylor series is equal \( f(x) \) in the interval;

where \( R \) is the radius of the convergence. The radius of the convergence for the Taylor series is determined by the following formula:
\[ R = \lim_{n \to \infty} \frac{1}{\frac{f^{(n+1)}(x)}{f^{(n)}(x)} \frac{1}{n+1}} \]  

\(|x - x_0| < R\) is called the interval of convergence. If \( R = \infty \), then the Taylor series converges for all \( x \); if \( R = 0 \), then the Taylor series converges only at \( x_0 \).

To approximate the derivative value as a function of the underlying variable using a Taylor series, we choose \( S^* \) to be \( x_0 \). The Taylor series of the derivative value is expressed as

\[
C(S^*) + C'(S^*)(S - S^*) + \frac{C''(S^*)}{2!}(S - S^*)^2 + \cdots + \frac{C^{(n)}(S^*)}{n!}(S - S^*)^n + \cdots
\]

\[ = \sum_{n=0}^{\infty} \frac{C^{(n)}(S^*)}{n!}(S - S^*)^n. \]  

It is reasonable to assume that the derivative \( C(S) \) is infinitely differentiable at \( S^* \). Even though the true derivative value is not infinitely differentiable at \( S^* \), the numerical analysis allows a Taylor series approximation of the derivative value to be infinitely differentiable at \( S^* \), meanwhile yielding an extremely close proxy to the true derivative value at \( S^* \). Thus, condition (1) is satisfied.

In order to determine the radius of the convergence \( R \), consider the partial differential equation of the derivative value

\[
C_t + rSC_S + \frac{\sigma^2}{2}S^2C_{SS} = rC
\]  

where \( C_t \) is the first order derivative of \( C \) with respect to \( t \), and \( C_{SS} \) is the second order derivative of \( C \) with respect to \( S \).

Taking derivative of the equation (5) with respect to \( S \), we have
\[
(r + \sigma^2)C_{ss} + \frac{\sigma^2}{2} S^2C_{sss} = 0 \tag{6}
\]

where \( C_{sss} \) is the third order derivative of \( C \) with respect to \( S \). Denote \( C^{(n)}(S) \) as the \( n \)-th order derivative of \( C \) with respect to \( S \). Taking derivative of equation (6) with respect to \( S \), we have

\[
(r + \frac{3}{2} \sigma^2)C^{(3)}(S) + \frac{\sigma^2}{2} SC^{(4)}(S) = 0.
\]

Continuously taking derivatives with respect to \( S \), we have

\[
\left( r + \frac{n}{2} \sigma^2 \right) C^{(n)}(S) + \frac{\sigma^2}{2} SC^{(n+1)}(S) = 0.
\]

Thus,

\[
\frac{C^{(n+1)}(S)}{C^{(n)}(S)} = \frac{\sigma^2 / 2}{r + n\sigma^2 / 2} \tag{7}
\]

Substituting (7) into (3) produces the radius of convergence

\[
R = \frac{1}{\lim_{n \to \infty} \left| \frac{C^{(n+1)}(S^*)}{C^{(n)}(S^*)} \right|^n} = \frac{1}{\lim_{n \to \infty} \left( \frac{\sigma^2 / 2}{r + n\sigma^2 / 2} \right)^n} = S^*. \]

Hence, the interval of convergence of the derivative value is

\[
|S - S^*| < S^*,
\]

or

\[
0 < S < 2S^*.
\]

Therefore, if \( S^* \) is sufficiently large, the Taylor series of the derivative value would converge in the interval of interest. Condition (2) is satisfied.
Since the underlying variable is a real variable in contrast to a complex variable, condition (3) is satisfied.

All three conditions being met, the Taylor series (4) converges to $C(S)$ in the interval $(0, 2S^*)$. In other words, the Taylor series (4) represents the derivative value as a function of the underlying variable over the interval $(0, 2S^*)$, when $S^*$ is sufficiently large. For instance, $S^*$ can be chosen to be the mean of the future underlying asset values.

When transforming the Taylor series of the derivative value (4) into polynomials, including more polynomials produces a more precise approximation of the derivative value. Nevertheless, Longstaff and Schwartz (1999) suggest that using more than three basis functions do not change the numerical results. When there are more than one state variable, the set of basis functions need only grow polynomially rather than exponentially. Therefore, for a derivative that is a function of $t$ and the underlying value only at $t$, the following regression function is employed,

$$C_t = \sum_{k=0}^{3} \beta_i, s_t^k + u_t.$$ 

In contrast, for a derivative that is a function of $t$, and the underlying value at $t$, as well as some function of the path followed by the asset price, $g(t, S)$, the following regression function is employed:

$$C_t = \beta_1 + \beta_2 S_t + \beta_3 S_t^2 + \beta_4 S_t^3 + \beta_5 g_t + \beta_6 g_t^2 + \beta_7 g_t^3 + \beta_8 S_t g_t + \beta_9 S_t^2 g_t + \beta_{10} S_t^2 g_t + u_t.$$
III. Example

Consider a 5-year Asian put option for which the payoff depends on the arithmetic average of the stock price during the life of the option, with a strike of $50. Assume the stock price at initiation is $50 and the stock price follows a log-normal diffusion process,

\[ dS = \mu S dt + \sigma S dz, \]

where \( \mu \) is the expected instantaneous return on the stock and \( \sigma \) is the volatility of the stock. Assume the volatility is 20% and the risk-free interest rate is 7%. To derive its distribution for year 2 after initiation, firstly simulate 10,000 underlying stock price paths over five years, starting with the initial stock price of $50. These simulations provide us with the underlying stock price under 10,000 different market scenarios over different periods of time. Secondly, calculate the ex-post realized payoff of each Asian put option at year 2 in each path by discounting the payoff along the path back to year 2. The payoff at maturity of the option is \( \max(S_{\text{ave,5}}^i - K, 0) \), where \( S_{\text{ave,5}}^i \) represents the arithmetic average stock value over the life of the option in path \( i \). Discounting this payoff back to year 2 yields the ex-post realized payoff \( e^{-0.07(t-2)} \max(S_{\text{ave,5}}^i - K, 0) \). This produces a sample of Asian put option values under different market scenarios. Thirdly, run a least squares regression of these option values on functions of the underlying stock price at year 2, \( S_2^i \), as well as the arithmetic average stock price between year 0 and year 2, \( S_{\text{ave,2}}^i \), as follows:

\[
C_2^i = \beta_1 + \beta_2 S_2^i + \beta_3 \left(S_2^i\right)^2 + \beta_4 \left(S_2^i\right)^3 + \beta_5 S_{\text{ave,2}}^i + \beta_6 \left(S_{\text{ave,2}}^i\right)^2 + \beta_7 \left(S_{\text{ave,2}}^i\right)^3 + \beta_8 S_{\text{ave,2}}^i S_{\text{ave,2}}^i + \beta_9 \left(S_{\text{ave,2}}^i\right)^2 + \beta_{10} \left(S_{\text{ave,2}}^i\right)^2 \left(S_{\text{ave,2}}^i\right) + u_i, \quad u_i \sim N(0, \sigma_u^2), \quad i = 1, 2, \ldots, 10000\]
The estimated option values from the regression are the conditional distribution of the Asian put option value at year 2. Table 1 presents the distribution of the Asian put option at year 2 after initiation. It includes the expected value, the standard deviation and the different percentiles of the distribution.

<table>
<thead>
<tr>
<th></th>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distribution of the value of a 5-year Asian put option on a stock 2 years in the future.</td>
<td></td>
</tr>
<tr>
<td>Expected Value</td>
<td>1.7770</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.9998</td>
</tr>
<tr>
<td>1% percentile</td>
<td>-0.6986</td>
</tr>
<tr>
<td>5% percentile</td>
<td>-0.5626</td>
</tr>
<tr>
<td>10% percentile</td>
<td>-0.5103</td>
</tr>
<tr>
<td>25% percentile</td>
<td>-0.2794</td>
</tr>
<tr>
<td>50% percentile</td>
<td>0.5508</td>
</tr>
<tr>
<td>75% percentile</td>
<td>2.7488</td>
</tr>
<tr>
<td>90% percentile</td>
<td>6.0752</td>
</tr>
<tr>
<td>95% percentile</td>
<td>8.0978</td>
</tr>
<tr>
<td>99% percentile</td>
<td>12.8939</td>
</tr>
</tbody>
</table>

**IV. Testing**

In order to test the performance of LSM in terms of estimating the distribution of the derivative value at a future point in time, it is necessary to find a benchmark to which the estimated distribution can be compared. Take an example of a European option on a stock. The actual distribution of a European option on a stock is of course known. However, I will proceed as if I do not know what the correct relation between the options price and the stock price at a future time. By comparing the option distribution estimated by LSM method with the actual distribution obtained through the Black-Scholes formula, we are able to see how well the LSM method works.
The option in the following example is a 5-year put option on a $50 stock with a strike of $50. The stock price follows a log-normal diffusion process,

\[ dS = \mu S \, dt + \sigma S \, dz \]

where \( \mu \) is the expected instantaneous return on the stock and \( \sigma \) is the volatility of the stock. Assume the volatility is 20% and the risk-free interest rate is 7%. 10,000 simulations of the stock price are generated using 1/252 years as the time step for each simulation over 5 years. Consider the put option distribution in year 2 when the option has 3 years to maturity. The payoff of the put option at maturity is \( \max \left( K - S_i^5, 0 \right) \),

where \( S_i^5 \) is the simulated stock price in the \( i^{th} \) path at year 5. The payoff is then discounted back along the path to year 2. Denote it as \( C_i^2 \). This is the ex-post realized payoff of the option. I then run a least squares regression of these ex-post realized payoff values on functions of the underlying stock price of year 2, \( S_i^2 \). The regression function is a polynomial function of the underlying stock:

\[ C_i^2 = \sum_{j=0}^{3} \beta_j X_i^j + \varepsilon, \quad i = 1, 2, \ldots, 10,000 \]

where \( \beta_j \) are the regression parameters and \( \varepsilon \) follows a standard normal distribution, i.e. \( \varepsilon \sim N(0,1) \). Since the stock price follows a log-normal distribution, the explanatory variable \( X \) is the log value of the stock price, \( \log \left( S_i^2 \right) \). The estimated option values from the regression, \( \hat{C}_i^2 \), are then the conditional distribution of the option value at year 2.

Figure 1 shows the 10,000 values of the option at the 2-year point in the simulation when the option has 3 years to maturity. The horizontal axis shows the stock price that was simulated, ranging from $20 to $155, and the vertical axis shows the
corresponding option value. The cloud indicates the ex-post realized payoffs of the
option, $C_t$. The thick dotted curve in the figure plots the option value estimated by the
least-squares regression $\hat{C}_t$. The thin curve in the figure plots the option price calculated
by the Black-Scholes formula. The curves coincide except when the stock price is either
extremely low or extremely high.

Table 2 provides a comparison of the expected value, the standard deviation and
the different percentiles of the distribution estimated by LSM and the distribution
calculated by the Black-Scholes formula. The column Relative Difference indicates the
difference between the two values divided by the True value.

<table>
<thead>
<tr>
<th></th>
<th>True Distribution</th>
<th>LSM Distribution</th>
<th>Relative Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Value</td>
<td>2.6400</td>
<td>2.6551</td>
<td>0.0057</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.8402</td>
<td>2.9547</td>
<td>0.0403</td>
</tr>
<tr>
<td>1% percentile</td>
<td>0.0170</td>
<td>-0.0502</td>
<td>-3.9542</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.0926</td>
<td>-0.0230</td>
<td>-1.2485</td>
</tr>
<tr>
<td>10% percentile</td>
<td>0.1937</td>
<td>0.0611</td>
<td>-0.6845</td>
</tr>
<tr>
<td>25% percentile</td>
<td>0.5828</td>
<td>0.4962</td>
<td>-0.1484</td>
</tr>
<tr>
<td>50% percentile</td>
<td>1.6374</td>
<td>1.6975</td>
<td>0.0367</td>
</tr>
<tr>
<td>75% percentile</td>
<td>3.7217</td>
<td>3.8013</td>
<td>0.0214</td>
</tr>
<tr>
<td>90% percentile</td>
<td>6.6161</td>
<td>6.5996</td>
<td>-0.0025</td>
</tr>
<tr>
<td>95% percentile</td>
<td>8.6153</td>
<td>8.6107</td>
<td>-0.0005</td>
</tr>
<tr>
<td>99% percentile</td>
<td>12.8574</td>
<td>13.4289</td>
<td>0.0444</td>
</tr>
</tbody>
</table>

Both Figure 1 and Table 2 indicate that the distribution estimated by LSM is very
close to the true distribution.
One issue that is worth noting here is that all the simulated paths should be utilized to regress the option value on functions of the underlying variable. Figure 2 is similar to Figure 1 except that the regression is based on only in-the-money paths. The cloud indicates the ex-post realized payoffs of the option only when they are in-the-money. The thick dotted curve in the figure shows the option value estimated by LSM with only in-the-money paths. The thin curve in the figure shows the option value derived by Black-Scholes formula. Clearly, the regression with only in-the-money paths provides an upward biased distribution.

Table 3 provides the expected value, the standard deviation and different percentile data for both distributions.

<table>
<thead>
<tr>
<th></th>
<th>True Distribution</th>
<th>LSM Distribution</th>
<th>Relative Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Value</td>
<td>4.8528</td>
<td>8.9543</td>
<td>0.8452</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>3.5117</td>
<td>2.4559</td>
<td>-0.3007</td>
</tr>
<tr>
<td>1% percentile</td>
<td>0.2202</td>
<td>4.9190</td>
<td>21.3418</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.6441</td>
<td>5.5500</td>
<td>7.6168</td>
</tr>
<tr>
<td>10% percentile</td>
<td>1.0485</td>
<td>6.0717</td>
<td>4.7907</td>
</tr>
<tr>
<td>25% percentile</td>
<td>2.1684</td>
<td>7.1969</td>
<td>2.3191</td>
</tr>
<tr>
<td>50% percentile</td>
<td>4.0561</td>
<td>8.6036</td>
<td>1.1211</td>
</tr>
<tr>
<td>75% percentile</td>
<td>6.7482</td>
<td>10.3339</td>
<td>0.5314</td>
</tr>
<tr>
<td>90% percentile</td>
<td>9.7441</td>
<td>12.2099</td>
<td>0.2531</td>
</tr>
<tr>
<td>95% percentile</td>
<td>11.7509</td>
<td>13.4923</td>
<td>0.1482</td>
</tr>
<tr>
<td>99% percentile</td>
<td>15.4701</td>
<td>15.9979</td>
<td>0.0341</td>
</tr>
</tbody>
</table>
Both Figure 2 and Table 3 indicate that there is an upward bias in the results derived through LSM methodology when using only in-the-money paths to run the regression.

The previous two examples indicate that LSM provides an efficient and effective methodology in generating the conditional distribution of a derivative value when the underlying variable follows a log-normal distribution. The following example examines the performance of LSM when the underlying variable follows a mean-reverting process.

Jamshidian (1989) derives a closed-form solution for a European option on pure discount bonds, assuming a mean-reverting Gaussian interest rate model as in Vasicek (1977). The diffusion process for the underlying interest rate is

$$dr = a(b - r)dt + \sigma dz,$$

where $a$ is the mean-reversion speed of the interest rate, $b$ is the mean-reversion level, and $\sigma$ is the volatility of the interest rate.

The value of the European call option on a pure discount bond is

$$LP(0,s)N(h) - XP(0,T)N(h - \sigma_p),$$

where $L$ is the bond principal, $s$ is the bond maturity, $X$ is the strike price, and $T$ is the option maturity,

$$h = \frac{1}{\sigma_p} \ln \frac{LP(0,s)}{P(0,T)}X + \frac{\sigma_p}{2},$$

$$\sigma_p = \frac{\sigma}{a} \left[1 - e^{-a(t-T)}\right] \sqrt{\frac{1-e^{-2at}}{2a}},$$

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)},$$

with
The option in the following example is a 5-year put option on a 10-year pure discount bond that pays $1 at maturity. The strike price is $0.7. Assume the mean-reversion speed of the short rate is 0.05 and the mean-reversion level is 0.1, and the volatility of the short rate is 0.01. 10,000 simulations of the interest rate are generated using 1/252 years as the time step for each simulation over 5 years. Let us take a look at the put option distribution in year 2 when the option has 3 years to maturity. The payoff of the put option is computed at the end of each interest rate path by $\max\left(K - P_5^t, 0\right)$, where $P_5^t$ is the pure discount bond price at year 5. $P_5^t$ is computed by plugging the simulated interest rate $r_5^t$ into equation (8c). The option payoff is then discounted back along the path to year 2. Denote it as $C_1^t$. This is the ex-post realized payoff of the option. I then run a least squares regression of these ex-post realized payoff values on functions of the underlying interest rate for year 2, $r_2^t$. The regression function is a polynomial function of the underlying rates:

$$C_2^t = \sum_{i=0}^{10000} \beta_i r_2^t + \varepsilon, \quad i = 1, 2, \ldots, 10,000,$$

where $\beta_i$ represent the regression parameters and $\varepsilon \sim N(0,1)$. The estimated option values from the regression, $\hat{C}_2^t, \quad i = 1, 2, \ldots, n$, are then the conditional distribution of the option value at year 2.
Figure 3 shows the 10,000 values of the option at the 2-year point in the simulation when the option has 3 years to maturity. The horizontal axis shows the interest rate that was simulated, ranging from 0 to 0.105, and the vertical axis shows the corresponding option value. The cloud indicates the ex-post realized payoffs of the option, $C^t$. The thick dotted curve in the figure plots the option value estimated by the least-squares regression $\hat{C}^t$. The thin curve in the figure plots the option price calculated by the Jamshidian formula, equation (8). The curves coincide except when the interest rate is extremely low.

Table 4 provides the comparison of the expected value, the standard deviation and the different percentiles of the distribution estimated by LSM and the distribution calculated by formula (8). The column Relative Difference indicates the difference between the two values divided by the True value.
Table 4
Comparison of the distribution of the price of a 5-Year put option on a pure discount bond 2 years in the future. The true distribution is based on the Jamshidian model. The relative difference is the LSM value less the true value as a proportion of the true value.

<table>
<thead>
<tr>
<th></th>
<th>True Distribution</th>
<th>LSM Distribution</th>
<th>Relative Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Expected Value</strong></td>
<td>0.0345</td>
<td>0.0351</td>
<td>0.0173</td>
</tr>
<tr>
<td><strong>Standard Deviation</strong></td>
<td>0.0240</td>
<td>0.0243</td>
<td>0.0128</td>
</tr>
<tr>
<td>1% percentile</td>
<td>0.0021</td>
<td>0.0024</td>
<td>0.1466</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.0052</td>
<td>0.0047</td>
<td>-0.0893</td>
</tr>
<tr>
<td>10% percentile</td>
<td>0.0085</td>
<td>0.0083</td>
<td>-0.0290</td>
</tr>
<tr>
<td>25% percentile</td>
<td>0.0164</td>
<td>0.0168</td>
<td>0.0276</td>
</tr>
<tr>
<td>50% percentile</td>
<td>0.0293</td>
<td>0.0302</td>
<td>0.0304</td>
</tr>
<tr>
<td>75% percentile</td>
<td>0.0474</td>
<td>0.0482</td>
<td>0.0166</td>
</tr>
<tr>
<td>90% percentile</td>
<td>0.0676</td>
<td>0.0682</td>
<td>0.0090</td>
</tr>
<tr>
<td>95% percentile</td>
<td>0.0815</td>
<td>0.0822</td>
<td>0.0086</td>
</tr>
<tr>
<td>99% percentile</td>
<td>0.1080</td>
<td>0.1096</td>
<td>0.0146</td>
</tr>
</tbody>
</table>

Both Figure 3 and Table 4 indicate that the distribution estimated by LSM is very close to the true distribution.

Finally, consider an example of an exotic option, specifically a barrier option, to examine the performance of the LSM methodology. A down-and-out option is a regular call option that ceases to exist when the underlying asset price reaches a certain barrier level. When the barrier level is below the strike price, the value of a down-and-out option is

$$C_{do} = C_{BS} - C_{di}$$  \hspace{1cm} (10)

where $C_{BS}$ is the value of the corresponding regular European call option, and $C_{di}$ is a down-and-in option where
The down-and-out option in the following example is a 5-year option on a $50 stock with a strike price of $60 and a barrier level of $45. The stock price follows a log-normal diffusion process,

\[ dS = \mu Sdt + \sigma Sdz, \]

where $\mu$ is the expected instantaneous return on the stock and $\sigma$ is the volatility of the stock. Assume the volatility is 20% and the risk-free interest rate is 7%. 10,000 simulations of the stock price are generated using 1/252 year as the time step for each simulation over 5 years. Consider the barrier option distribution in year 2 when the option has 3 years to maturity. When a whole path of the simulated stock price is above the barrier, the option payoff at maturity is \( \max(K - S^i_t, 0) \), where \( S^i_t \) is the simulated stock price in that path \( i \) at maturity. The payoff is then discounted back to year 2. Denote it as \( C^i_2 \). When any point in a simulated stock path hits or falls below the barrier $45, the option payoff of this path is zero. The value of the corresponding underlying stock at year 2, \( S^i_2 \), depends on the point where the path hits the barrier. If the path hits the barrier after year 2, \( S^i_2 \) is the simulated stock value in path \( i \) at year 2; if the path hits the barrier before year 2, \( S^i_2 \) is set to be at the barrier.

\[
C_{di} = S_i \left( \frac{H}{S_i} \right)^{2\lambda} N(y) - X e^{-\tau} \left( \frac{H}{S_i} \right)^{2\lambda-2} N \left( y - \sigma \sqrt{T-t} \right), \quad (10a)
\]

\[
\lambda = \frac{r + \sigma^2/2}{\sigma^2}, \quad (10b)
\]

and

\[
y = \frac{\ln \left[ \frac{H^2}{(S_i X)} \right]}{\sigma \sqrt{T-t}} + \lambda \sigma \sqrt{T-t}. \quad (10c)
\]
Figure 4 shows the 10,000 values of the barrier option at the 2-year point in the simulation when the option has 3 years to maturity. The horizontal axis shows the stock price that was simulated, ranging from $45 to $150, and the vertical axis shows the corresponding option value. The cloud indicates the ex-post realized payoffs of the option, \( C_t \). The thick dotted curve in the figure plots the option value estimated by the least-squares regression \( \hat{C}_t \). The thin curve in the figure plots the option price calculated by equation (10). The curves coincide except when the stock price is extremely high.

Table 5 provides the comparison of the expected value, the standard deviation and the different percentiles of the distribution estimated by LSM and the distribution calculated by the closed-form formula. The column Relative Difference indicates the difference of the two values divided by the True value.

<table>
<thead>
<tr>
<th></th>
<th>True Distribution</th>
<th>LSM Distribution</th>
<th>Relative Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Value</td>
<td>8.9942</td>
<td>9.1427</td>
<td>0.0165</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>14.0554</td>
<td>14.4189</td>
<td>0.0259</td>
</tr>
<tr>
<td>1% percentile</td>
<td>0.0000</td>
<td>0.0553</td>
<td>--</td>
</tr>
<tr>
<td>5% percentile</td>
<td>0.0000</td>
<td>0.0553</td>
<td>--</td>
</tr>
<tr>
<td>10% percentile</td>
<td>0.0000</td>
<td>0.0553</td>
<td>--</td>
</tr>
<tr>
<td>25% percentile</td>
<td>0.0000</td>
<td>0.0553</td>
<td>--</td>
</tr>
<tr>
<td>50% percentile</td>
<td>0.0000</td>
<td>0.0553</td>
<td>--</td>
</tr>
<tr>
<td>75% percentile</td>
<td>15.6088</td>
<td>15.8056</td>
<td>0.0126</td>
</tr>
<tr>
<td>90% percentile</td>
<td>29.4912</td>
<td>29.6304</td>
<td>0.0047</td>
</tr>
<tr>
<td>95% percentile</td>
<td>38.3214</td>
<td>38.6051</td>
<td>0.0074</td>
</tr>
<tr>
<td>99% percentile</td>
<td>57.2326</td>
<td>59.5163</td>
<td>0.0399</td>
</tr>
</tbody>
</table>
Both Figure 4 and Table 5 indicate that the distribution estimated by LSM is very close to the true distribution.

**Risk Neutral World vs. Real World**

Risk Management Analysis concerns the distribution of future losses that could be experienced in the real world. The simulation of the underlying variable up to the future time of interest should consequently be in the real world. This real world distribution of the future underlying variable in return results in the real world distribution of the future derivative value. The underlying variable simulated from the future time of interest to maturity, in contrast, should be in the risk neutral world. Since all the derivative pricing models are developed in the risk neutral valuation framework, this risk neutral world simulation facilitates the correct valuation of the derivative at the future time of interest.

**IV. Conclusion**

The distribution of a derivative value at a future time has been a main focus for risk management. In recent years, there has been considerable interest in the derivation of the distribution of exotic options when a closed-form solution to the derivative valuation is not available. This paper employs the least-squares-Monte-Carlo (LSM) methodology to derive the distribution at a future time of path dependent products and derivatives with American-style exercise features.

In contrast to the standard Monte-Carlo simulation procedure, the LSM method combines the least-squares regression technique with the Monte-Carlo simulation to derive the conditional distribution of some exotic derivatives by regressing the ex-post realized derivative values on functions of the underlying values. It circumvents the
computational difficulty of the pure Monte-Carlo procedure when deriving the expected future derivative value given certain future underlying values. Instead of calculation-intensive sub-Monte-Carlo simulations, the implementation of the LSM requires only least-squares regressions.

LSM methodology makes it far easier and more computationally efficient for those managing risk to calculate derivative distributions for a large number of exotic options that do not possess closed-form solutions. By means of several examples, using options with closed-form solutions, this paper demonstrates the ability of LSM to produce excellent estimates of derivative distribution at a reasonable computational cost.
Reference


This figure presents the comparison of a stock put option distribution estimated by using LSM method (the dotted thick curve) with that derived by using Black-Scholes formula (the thin curve). The horizontal axis shows the stock price and the vertical axis shows the option value.
Figure 2. Distribution of Put Option on Stock
(LSM of in-the-money paths only vs Analytical Solution)

This figure presents the comparison of a stock put option distribution estimated by using LSM method when only in-the-money paths are used (the dotted thick curve) with that derived by using Black-Scholes formula (the thin curve). The horizontal axis shows the stock price and the vertical axis shows the option value.
Figure 3. Distribution of Call Option on Pure Discount Bond 
(LSM vs. Analytical Solution)

This figure presents the comparison of a pure discount bond call option distribution estimated by using LSM method (the dotted thick curve) with that derived by using Jamshidian formula (the thin curve). The horizontal axis shows the interest rate and the vertical axis shows the option value.
Figure 4. Distribution of Barrier Option on Stock
(LSM vs. Analytical Solution)

This figure presents the comparison of a stock down-and-out option distribution estimated by using LSM method (the dotted thick curve) with that derived by using the analytical formula (the thin curve). The horizontal axis shows the stock price and the vertical axis shows the option value.
Comparison of Two Credit Risk Portfolio Models: CreditMetrics and CreditRisk+
I. Introduction

For risk management purposes, financial institutions are interested in calculating a complete probability distribution for the credit losses that could be experienced in the real world. This distribution can be used to determine the institution's credit VaR - a credit value at risk measure that indicates the level of credit loss that we are x % certain will not be exceeded at a given time T. In addition, the distribution can be used to determine the level of economic capital required to cover the risk of unexpected credit losses.

There are two types of credit risks: credit rating change risk and credit default risk. Credit rating change risk is the risk of financial loss owing to changes in credit rating. For example, if downgrading from A to BBB occurred, all outstanding contracts with the counterparty would need to be revalued, and a "mark-to-market" credit loss would be experienced. Credit default risk is the risk that a counterparty is unable to meet its financial obligations. When a counterparty defaults, the financial institution incurs a loss equal to the amount owed by the counterparty less any recovery amount received as a result of foreclosure, liquidation or restructuring of the defaulted counterparty.

The main task of risk management is to analyze the uncertainty associated with the value of the assets in the portfolio. Measuring the uncertainty or variability of losses in a portfolio is crucial to the effective management of credit risk. The expected loss itself is easily calculated and accounted for in the standard pricing model, which takes account of market risk. Thus, sufficient earnings should be generated through adequate pricing to absorb any expected loss. However, the actual credit loss suffered in any one period could be significantly more than the expected loss. Therefore, the financial institution
needs to put aside an amount of money called economic capital to ensure continued solvency in the event of large unexpected credit losses.

This paper compares two of the most popular credit risk portfolio models: CreditMetrics by J.P. Morgan and CreditRisk+ by Credit Suisse First Boston. The CreditMetrics approach is based on Merton’s model of a firm’s capital structure: a firm defaults when its asset value falls below its liabilities. A counterparty’s default probability therefore depends on the amount by which assets exceed liabilities, and the volatility of the assets. If changes in asset value are normally distributed, the default probability can be expressed as the probability of a standard normal variable falling below some critical value. However, CreditMetrics not only considers the value changes due to default, but also those due to credit rating changes. Extending the intuition for the default case, CreditMetrics assumes there are several critical or “threshold” values that are compared with the actual value of a firm’s assets to determine its credit rating at the end of each period. Therefore, we only need to model the company’s change in asset value in order to describe its credit rating evolution. Credit rating changes of different counterparties in the portfolio are related because changes in the asset value of those counterparties may all in turn be related to certain common economic factors. Thus, the pairwise asset correlations are the data input used to determine joint credit rating change behavior. This methodology requires the portfolio loss distribution to be calculated using Monte Carlo simulations.

CreditRisk+ utilizes ideas from the insurance industry to model the loss distribution. In contrast to CreditMetrics, which takes account of both credit rating change risk and default risk, CreditRisk+ examines only the default component of
portfolio credit risk. Counterparties are allocated to different "sectors", each of which has a mean default rate and a default rate volatility. The default rate is assumed to follow the Gamma distribution, and the conditional default distribution is assumed to follow the Poisson distribution. The convolution of the Poisson distribution with the Gamma distribution gives rise to an unconditional default distribution that can be identified as a Negative Binomial distribution. Once the random recovery rate is also taken into account, an analytic solution for the distribution of portfolio losses may be obtained by utilizing the specific properties of the probability generating function. Default rate volatility and sector analysis rather than default correlations are used to model the joint-default behavior of counterparties.

Both approaches have their advantages and their drawbacks. CreditMetrics focuses on quantifying the probability distribution of the losses arising not only from default but also from credit rating changes. However, the need for Monte Carlo simulations to implement the technique is somewhat burdensome. CreditRisk+ arrives at an analytical solution, which is easier to implement than the CreditMetrics approach. However, the CreditRisk+ approach focuses only on quantifying the probability distribution of losses arising from counterparty default, completely ignoring the possibility of credit rating changes.

This paper applies both approaches to a harmonized set of input parameters. The results suggest that the two approaches may lead to quite different conclusions for a financial institution that is using the models to determine how much economic capital to hold.
II. Model Methodologies

CreditMetrics

Merton's model assumes that the corporation has two types of claims. One is a single and homogenous debt claim; the other is an idealized residual equity claim. The firm promises to pay a fixed amount of dollars, \( F \), to the bondholders on a specific date, \( T \). If the value of the firm's assets at time \( T, V_T \), is not sufficient to pay the lenders this amount, the bondholders take over the company and get paid \( V_T \) and shareholders receive nothing since the residual value is zero. When \( V_T \) is greater than \( F \), the bondholders get paid the promised amount \( F \), and the value of the equity at time \( T \) is the residual amount \( V_T - F \). Thus the payoff to the shareholders at time \( T \) is

\[
E(V, T) = \max(0, V_T - F),
\]

And the payoff to the bondholders at time \( T \) is

\[
D(V, T) = \min(V_T, F) = F - \max(0, F - V_T).
\]

Therefore, we can consider the equity as a call option on the firm's asset value and the debt as a portfolio of a risk-free bond with face value \( F \) and a put option on the firm's asset value.

By analogy with the Black-Scholes formula, we can then calculate the current equity value:

\[
E(V, 0) = V_0 N(d_1) - F e^{-rT} N(d_2) \tag{2.1}
\]

\[
d_1 = \frac{\ln(V_0/F) + (r + \frac{\sigma^2}{2})T}{\sigma \sqrt{T}}
\]

\[
d_2 = d_1 - \sigma \sqrt{T}
\]
When the firm's asset value is not high enough to repay the promised amount to the bondholders, that is, when \( V_T < F \), the bondholders take over the firm, and default occurs. Using risk-neutral valuation, we can express the current equity value as follows,

\[
E(V, 0) = \Pr(\text{no default}) \times (V_T - F) \times e^{-rT} + \Pr(\text{default}) \times 0
\]

\[
= [1 - \Pr(\text{default})] \times V_T \times e^{-rT} - F \times e^{-rT} \times [1 - \Pr(\text{default})]. \tag{2.2}
\]

Comparing equation (2.1) with (2.2), we have

\[
\Pr(\text{default}) = N(-d_2).
\]

Therefore, the risk-neutral probability of default in Merton's model is \( N(-d_2) \),

where

\[
N(-d_2) = \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx
\]

and

\[
d_2 = \frac{\ln(V_0/F) + (r - \sigma^2/2)T}{\sigma \sqrt{T}}.
\]

Rating agencies such as Moodys and Standard and Poors provide credit ratings for firms. A credit rating is a measure of the credit quality of a firm and the possibility of a change in this credit quality is a fundamental source of credit risk. Typically, rating agencies divide firms into different categories according to their quality, and indicate the assigned category with an alphabetic or numeric label. If the rating system is correct the rating gives an ordinal ranking of the default probabilities across the categories.

Quantitative frameworks, such as the CreditMetrics or CreditRisk+ models, give meaning to the rating labels by linking each one with a default probability. Thus credit quality is generally used to refer to the relative chance of default, which is measured by credit rating. Since CreditMetrics also considers the possibility of credit rating upgrades and downgrades as a source of credit risk, they utilize credit rating as an indication of both credit...
the chance of default and the chance of credit rating changes, which are together referred to as credit rating migrations.

In practice, credit rating migrations are not readily observable. So we need to model these migrations through a process which we understand and can observe. Within the context of Merton's model, CreditMetrics proposes that a firm's asset value be the basis for predicting its credit rating migration.

For ease of discussion, let's first consider default risk exclusively. In Merton's model, the value of a firm's assets determines its ability to pay its debt holders. We may assume that if the firm's asset value falls below a specific level, the firm will not be able to fully pay back its liabilities outstanding and will therefore default. That specific level is termed the default threshold by CreditMetrics. Thus, given the underlying firm value, which is random with some known distribution, we can calculate the probability that the value of assets at the end of the period in question falls below the default threshold. The probability of default is indicated in Figure 1.

Up to this point, we have described the default process for each counterparty individually according to its asset value process. Joint default events among counterparties in the portfolio are related because the counterparties' changes in asset value are correlated. Asset correlation measures the degree to which two counterparties' asset values tend to move in common. If the asset values of two counterparties are highly positively correlated, then if one asset value drops, the other will be very much likely to drop as well. So if one counterparty defaults, there is a high chance that the other counterparty will default at the same time. Therefore, with highly positively correlated asset values, the probability of both counterparties defaulting simultaneously is high. If
the asset values of two counterparties are not correlated, then when one asset value drops, the other might drop or increase or stay unchanged since the change in asset value of the latter is independent of the change in asset value of the former. Therefore the probability of simultaneous default is low compared to the previous case in which both asset values are positively correlated. For example, consider two counterparties, A and B, whose asset returns are normally distributed and correlated, with correlation coefficient \( \rho_{AB} \) \( (\rho_{AB} > 0) \), and

\[
\Delta V_A \sim N(0, \sigma_A^2), \\
\Delta V_B \sim N(0, \sigma_B^2).
\]

Probability of counterparty A's default is \( P_A \) and that of counterparty B is \( P_B \). Since we know the distribution of asset values, we can use these probabilities to calculate the default threshold values \( Z_{D} \) and \( Z_{D}' \) for counterparty A and B respectively.

The covariance matrix for the bivariate normal distribution is

\[
\Sigma = \begin{bmatrix}
\sigma_A^2 & \rho_{AB} \sigma_A \sigma_B \\
\rho_{AB} \sigma_A \sigma_B & \sigma_B^2
\end{bmatrix}
\]

To compute the probability that both counterparties default, which is the joint probability that the asset return for counterparty A falls below \( \Delta Z_{D} \) while at the same time the asset return for counterparty B falls below \( \Delta Z_{D}' \), we have

\[
P_{AB} = \Pr(\Delta V_A < \Delta Z_D \text{ and } \Delta V_B < \Delta Z_D') = \int_{-\infty}^{Z_D'} \int_{-\infty}^{Z_D} f(\Delta V_A, \Delta V_B, \Sigma) \ d\Delta V_A \ d\Delta V_B
\]

where \( f(\Delta V_A, \Delta V_B, \Sigma) \) is the density function for the bivariate normal distribution with covariance matrix \( \Sigma \). With the distribution of both counterparties’ asset values given, the asset correlation between A and B determines their joint probability of default.
Figure 2 indicates the joint probability of default of counterparty A and B with a positive pairwise correlation.

In addition to default risk, CreditMetrics extends Merton’s model to include credit rating change risk. By assuming a counterparty’s credit rating is determined by its asset value, they come up with a series of credit rating upgrade or downgrade thresholds. The firm’s asset value relative to these thresholds determines its future rating.

A firm’s future asset value is assumed to be random with some known probability distribution. Consider the firm in Figure 3, which is currently rated BBB. If the value of assets at the end of the period in question decreases to the amount between $Z_D$ and $Z_{CCC}$, then the counterparty is downgraded to CCC. If the value of assets increases to the amount between $Z_A$ and $Z_{AA}$, then the firm is upgraded to AA. The probabilities of credit rating migration are indicated in Figure 3. The transition matrices in the “CreditMetrics – Technical document. The Benchmark for Understanding Credit Risk.” J.P. Morgan (1997) provide the data of these probabilities. Rating agencies such as moodys and S&P also regularly publish rating transition matrices that give these rating transition probabilities.
This way of modeling rating transitions can be extended to multiple firms in exactly the same way that the procedure for modeling defaults was extended from one firm to multiple firms. The asset values of the firms are assumed to be jointly normally distributed with some known correlation. This joint probability distribution is then overlaid with a grid. The horizontal divisions of the grid correspond to the rating transition points for one firm analogous to the values $Z_D, Z_{CCC}$, etc., in Figure 3. The vertical divisions in the grid correspond to the rating transition points for the other firm. Every cell in the grid thus corresponds to the case in which the first firm takes on one rating and the second firm takes on a second rating. The probability of firm A ending up rated A while firm B is rated CCC is

$$
\Pr\{A_{Z_{BBB}} < \Delta V_A < A_{Z_A} \text{ and } B_{Z_D} < \Delta V_B < B_{Z_{CCC}}\} = \int_{A_{Z_{BBB}}}^{A_{Z_A}} \int_{B_{Z_D}}^{B_{Z_{CCC}}} f(\Delta V_A, \Delta V_B, \Sigma) d\Delta V_B d\Delta V_A
$$

where $f(\Delta V_A, \Delta V_B, \Sigma)$ is the density function for the bivariate normal distribution for the asset values.

**CreditRisk+**

CreditRisk+ makes use of mathematical techniques common in loss distribution modeling in the insurance industry. There is no assumption about the causes of default. In contrast to the CreditMetrics, which takes account of both credit rating change risk and default risk, it only examines the default component of portfolio credit risk. The joint-default behavior of counterparties is incorporated by utilizing the default rate and standard deviation of each counterparty. Adding the individual rates yields a default rate for the entire portfolio. Similarly, adding the individual standard deviations yields the standard deviation of the default rate for the entire portfolio. CreditRisk+ models the
joint-default behavior of the counterparties by employing this standard deviation rather than by using default correlation as a direct input.

To obtain the distribution of default loss, CreditRisk+ first of all studies the distribution of default events involving one element of randomness, the default rate. Based on the distribution of the number of defaults, and incorporating a second element of randomness, the recovery rate, CreditRisk+ derives the distribution of default loss.

Since CreditRisk+ examines only default risk, there are only two possible states at the end of the period: default and no-default. If we assume that each counterparty has the same default rate $p$, it is natural to assume that among $N$ counterparties, the probability of $n$ defaults follows a binomial distribution, of which the probability density function is

$$
\Pr(n \ \text{defaults}) = \frac{N!}{n! (N-n)!} p^n (1-p)^{N-n}, \quad (2.3)
$$

When $p$ is small and $N$ is large, a binomial distribution can be approximated by a Poisson distribution. The probability density function of Poisson is

$$
\Pr(n \ \text{defaults}) = \frac{e^{-\mu} \mu^n}{n!}, \quad (2.4)
$$

where $\mu = Np$ is the sum of the default probabilities of all the counterparties. For $N$ counterparties that have different but fixed default rates, $p_1, p_2, \ldots, p_N$, equation (2.4) can be used to calculate the probability of $n$ defaults if $\mu = \sum_{j=1}^{N} p_j$, as long as there is a large number of counterparties, each with a small probability of default. This is usually the case.

The procedure just described assumes that default rates are constant. However, CreditRisk+ works in a world where the default rates for each counterparty are not fixed,
but stochastic. CreditRisk+ derives a closed-form solution that applies to this stochastic case, identifying the appropriate default distribution as one closely resembling a Negative Binomial Distribution:

\[ P(n \text{ defaults}) = (1 - p_k)^{a_k} \binom{n + \alpha_k - 1}{n} p_k^n, \]  

(2.5)

where \( \alpha_k = \frac{\mu_k}{\sigma_k^2} \), \( \beta_k = \frac{\sigma_k^2}{\mu_k} \), and \( p_k = \frac{\beta_k}{1 + \beta_k} \). The parameters \( \alpha_k \) and \( \beta_k \) are explained in the paragraph below.

CreditRisk+ derives this closed-form solution by separating the counterparties into different groups or “sectors”, each of which has a mean default rate and a default rate volatility. The default rate \( x_k \) for the kth sector is assumed to follow a Gamma distribution, with parameters \( \alpha_k \) and \( \beta_k \):

\[ x_k \sim \Gamma(\alpha_k, \beta_k), \]

where \( \alpha_k = \frac{\mu_k}{\sigma_k^2} \), \( \beta_k = \frac{\sigma_k^2}{\mu_k} \), and \( \mu_k \) and \( \sigma_k \) are the mean and volatility of the default rate respectively.

Within each sector, the counterparties are then divided into different subgroups, called “bands”, where each band consists of counterparties having the same exposure size. A unit amount of exposure \( L \), which serves as a base currency, needs to be chosen such that the exposure and the expected loss of each counterparty are expressed as multiples of the unit. This significantly reduces the amount of data incorporated into the calculation, strengthening one of the advantages of this approach – calculation efficiency.

Within each band, CreditRisk+ assumes that the default distribution follows the Poisson distribution, and the joint-default behaviors are independent conditional on the fixed...
default rates. The convolution of the Poisson distribution within each band with the Gamma distribution between bands within the same sector gives rise to an unconditional default distribution that closely resembles the Negative Binomial distribution, as indicated in equation (2.5).

To get from this distribution of the default rate to the final goal of obtaining the distribution of default loss, we need to build the randomness of the recovery into the model. This random effect is best described mathematically through a probability generating function. The distribution of loss has a closed-form solution, obtained through utilizing the special properties of the probability generating function. The probability of loss can be solved analytically from the following recurrence equation:

\[
A_{n+1} = \frac{1}{b_n(n+1)} \left[ \sum_{j=0}^{\min(r,n)} a_j A_{n-j} - \sum_{j=0}^{\min(n-1,n-1)} b_{j+1} (n-j) A_{n-j} \right].
\]

The parameters \(a_i (i = 0,1,2,...,r)\) and \(b_j (j = 0,1,2,...,s)\) satisfy the polynomials

\[
A(z) = a_0 + \cdots a_r z^r
\]

\[
B(z) = b_0 + \cdots b_s z^s.
\]

The calculation of the parameters \(a_i\) and \(b_j\) can be obtained by setting the polynomials into the following relation:

\[
\frac{A(z)}{B(z)} = \sum_{k=1}^{n} \frac{\mu_k \phi_k^{(k)}}{1 - \frac{\sum_{j=1}^{m(k)} \phi_j^{(k)} z_j^{(k)-1}}{\phi_j^{(k)}}},
\]

where \(v_j^{(k)}\) is the \(k\)th sector’s exposure, and \(e_j^{(k)}\) is the \(k\)th sector’s expected loss. Both are expressed as multiples of the unit \(L\).
In conclusion, by making the distributional assumptions discussed above, and by utilizing the properties of the probability generating function, CreditRisk+ arrives at an analytical solution to the probability of losses.

**Default Rate Volatility and Asset Correlation**

Before moving on to illustrate the implementation of both models, I will spend some extra time emphasizing the connections and the differences between the two approaches used to model joint default events among counterparties. Understanding both the connections and the differences is necessary for constructing a consistent set of data parameters to compare the recommendations of the two approaches.

The default rate, being assigned to each counterparty, represents the likelihood of a default event occurring. Published statistics show that one-year default rates among a group of rated companies within a given industry vary significantly year by year. The number of defaults during an economic expansion can be significantly lower than during an economic recession.

Such yearly statistics may be considered as the realization of a random variable. The average default rate over many years is the expected value of that random variable. The variation in default rates arises from some common factors, such as the state of the economy that may influence a large number of counterparties in the same way. Thus the uncertainty of the default rate can be related to the underlying variability in a relatively small number of common factors. For instance, a downward trend in the state of the economy may cause most of the counterparties in a portfolio to default.
Incorporating the effects of the common factors into the specification of default rates causes the loss distribution to exhibit a significant skewness with fatter tails. Such a fat-tail distribution is consistent with the specific credit portfolio distribution.

CreditRisk+ utilizes the default rate volatility to build up the fat-tail distribution. They assume the default rate is a random variable with some known mean default rate and a default rate volatility.

In contrast to the procedure used by CreditRisk+, CreditMetrics utilizes the default correlation. If defaults were uncorrelated, we would expect the default rate to be fairly stable over time; if defaults were perfectly correlated, then we would expect that in some years no counterparties would default, while in other years all the counterparties would do so.

CreditMetrics models the unobservable process of defaults through the observable process of changes in firms' asset value. A counterparty's default is triggered by the change in its asset value, which is correlated with the change in another counterparty's asset value. According to the factor model, the change in asset value can be decomposed into several orthogonal factors that are common to the returns of all assets under consideration. The parameters of the factors, the factor loadings, can be derived from the asset correlations. Given the factor loadings, the conditional probability of default, contingent on the factor loadings, can be obtained. The variation in the probability of default can then be calculated.

Now let us return to the previous example where the change in asset value was assumed to be drawn from a normal distribution:

$$\Delta V_i \sim N(0, \sigma_i^2).$$
Applying the factor model, we can decompose the asset value into $k$ factors:

$$
\Delta V_j = b_{j1}X_1 + b_{j2}X_2 + \cdots + b_{jk}X_k + b_j \epsilon_j
$$

where $X_1, X_2, \ldots, X_k$ are the mean zero factors common to the returns of all assets under consideration; $b_{j1}, b_{j2}, \ldots, b_{jk}$ are the factor loadings; $\epsilon_j$ is a noise term with a parameter $b$; $X_1, X_2, \ldots, X_k$ and $\epsilon_j$ are independent and identical standard normal distributed, i.e. $X_1, X_2, \ldots, X_k, \epsilon_j \sim iid \ N(0,1)$ and $b = \sqrt{\sigma^2_j - \sum b_{jm}^2}$. Then, $\Delta V_j \sim N\left(0, \sigma^2_j\right)$ follows.

The unconditional probability of default is:

$$
\bar{p}_j = \Pr(\Delta V_j < Z_D) = \Pr\left(\frac{\Delta V_j}{\sigma_j} < \frac{Z_D}{\sigma_j}\right) = N\left(\frac{Z_D}{\sigma_j}\right),
$$

where $N(\cdot)$ is the cumulative distribution function of the standard normal distribution.

To derive the probability of default conditional on the common factors we start with the change of the asset value based on the factor model:

$$
\Delta V_j = b_{j1}X_1 + b_{j2}X_2 + \cdots + b_{jk}X_k + \sqrt{\sigma^2_j - \sum b_{jm}^2} \epsilon_j.
$$

From $\Delta V_j < Z_D$, we then have

$$
\sum_{m=1}^{k} b_{jm} X_m + \sqrt{\sigma^2_j - \sum_{m=1}^{k} b_{jm}^2} \epsilon_j < Z_D
$$

$$
\Rightarrow \epsilon_j < \frac{Z_D - \sum b_{jm} X_m}{\sqrt{\sigma^2_j - \sum_{m=1}^{k} b_{jm}^2}}.
$$

Therefore, the conditional probability of default is
\[ p_j \mid x = \Pr(\Delta V_j \mid x < Z_D) = \Pr \left( \frac{Z_D - \sum_{m=1}^{k} b_{jm} X_m}{\sqrt{\sigma_j^2 - \sum_{m=1}^{k} b_{jm}^2}} < \frac{e_j}{\sqrt{\sigma_j^2 - \sum_{m=1}^{k} b_{jm}^2}} \right) = N \left( \frac{Z_D - \sum_{m=1}^{k} b_{jm} X_m}{\sqrt{\sigma_j^2 - \sum_{m=1}^{k} b_{jm}^2}} \right) \]

The variance of the probability of default can then be calculated by

\[ s_j^2 = \int (p_j \mid x - \bar{p}_j)^2 f(X) \, dX. \]

Let \( \sum_{m=1}^{k} b_{jm} X_m = Y_j \), then

\[ Y_j \sim N \left( 0, \sum_{m=1}^{k} b_{jm}^2 \right). \]

The conditional probability of default can be written as

\[ p_j \mid x = N \left( \frac{Z_D - Y_j}{\sqrt{\sigma_j^2 - \sum_{m=1}^{k} b_{jm}^2}} \right) \]

Plugging this into the equation of variance of default probability, we get

\[ s_j^2 = \int \left[ N \left( \frac{Z_D - Y_j}{\sqrt{\sigma_j^2 - \sum_{m=1}^{k} b_{jm}^2}} \right) - N \left( \frac{Z_D}{\sigma_j} \right) \right]^2 \frac{1}{\sqrt{2\pi} \left( \sum_{m=1}^{k} b_{jm}^2 \right)^{1/2}} \exp \left( -\frac{Y_j^2}{2 \sum_{m=1}^{k} b_{jm}^2} \right) dY_j, \]

where \( s_j \) is the standard deviation of the default rate, the data input for CreditRisk+.

Since the factor loadings, \( b_{jm} \), \( m = 1 \ldots k \), can be calculated from the asset correlation data, the above derivation of the relation between asset correlation and default rate volatility illustrates the connection between the two models. CreditMetrics incorporates the effect of the common factors into the specification of default rates by using asset correlation; CreditRisk+ does so by using default rate volatility.
III. Implementation of Both Models

I use the example provided in the J.P.Morgan(1997) to implement both approaches. The portfolio contains twenty corporate bonds with different ratings, principals, coupon rates, and maturities. The data are presented in Table 1.

CreditMetrics

In this example, the time horizon is chosen to be one year. Given different forward interest rate curves for different credit ratings, we can calculate what the new asset value will be at the end of year 1 for every possible credit rating. For example, suppose asset $j$ is currently rated as BBB. At the end of year 1, if it is upgraded to AA, then the new value for asset $j$ can be calculated by discounting the coupon payments and the principal at the forward rates for credit rating AA. If it is downgraded to BB, then forward interest rates for rating BB should be used to get the new value. In the case of default, the new bond value is calculated by multiplying its face value by the assumed recovery rate. Rating transition matrices such as those produced by Moody's and S&P contain the probability of credit rating migrations. These matrices provide not only the likelihood of default but also the likelihood of migrating to any possible credit quality state.

When a Monte Carlo simulation is used to calculate the distribution of losses, it is necessary to get the new credit rating at the end of year 1. There are three steps involved.

Step #1: Calculate the threshold values corresponding to each counterparty's credit rating change and default probability. These correspond to the $Z_D$, $Z_{CCC}$, etc. in Figure 3.
Given the probability of default, it is possible to work backwards to get the threshold in asset value that corresponds to default. Assuming changes in asset value are normally distributed with zero mean and unit variance, the default probability is the probability of a standard normal variable falling below some critical value $Z_D$.

$$P_D' = N(Z_D).$$

Then,

$$Z_D = N^{-1}(P_D'),$$

and $Z_D$ is the threshold for default.

Similarly, the probability of downgrading to CCC is the probability of a standard normal variable falling between $Z_D$ and $Z_{CCC}$, that is

$$P_{ccc}' = N(Z_{ccc}) - N(Z_D).$$

Then,

$$Z_{ccc} = N^{-1}[P_{ccc}' + N(Z_D)]$$
$$= N^{-1}[P_{ccc}' + P_D'],$$

where $Z_{ccc}$ is the threshold for rating change to CCC. Proceeding in this way all the threshold values can be calculated from the probability of credit rating migrations.

Step #2: Draw random correlated standard normal variables representing the change in asset value for each counterparty.

Due to the common factors that affect the fortunes of some or all of the counterparties simultaneously, changes in asset values are correlated. The correlation structure assumed in this experiment is shown in Table 2. Note there are five groups of counterparties. Within each group, the pairwise correlations are relatively high, while the
pairwise correlations across groups are lower. This might be the case when the portfolio contains counterparties from five different industries or from 5 different countries.

Step #3: Compare the standardized changes in asset values to the threshold values to determine the credit rating migration for each of the counterparties.

Consider a single counterparty. If the change in the asset value falls between the threshold values $Z_{bb}$ and $Z_b$, then this counterparty's new credit rating at the end of year 1 is BB. If the change in the asset value falls lower than $Z_D$, then this counterparty defaults at the end of year 1 and so on. By comparing the simulated random variables to the threshold values, we obtain the one to one mapping from asset value changes to credit rating changes.

After obtaining the new credit rating for each counterparty and using the forward zero curve to calculate the new asset values, we can get the new portfolio value at the end of year 1 by summing up all the counterparties' asset values after credit rating migration. The portfolio loss is simply the new portfolio value minus the original portfolio value at time 0. Repeating such a procedure 10,000 times yields the distribution of portfolio loss.

To make the comparison of CreditMetrics and CreditRisk+ more interesting, we implement the CreditMetrics approach eliminating the risk of credit rating changes. One of the major differences between the two approaches is that CreditMetrics takes account of the credit rating change risk while CreditRisk+ does not. So, by eliminating the rating change risk in the CreditMetrics model, we can separate the loss distribution results obtained by J. P. Morgan (1997) into two categories: those from default risk and those from credit rating change risk.
If we only take account of default risk, instead of the eight rating states AAA, AA, A, BBB, BB, B, CCC, and Default, there are only two states the counterparty could land at the end of year 1, Default and No Default. Therefore, in step #1 described above, there is only one threshold value $Z_\sigma$ for each counterparty. And in step #3, if the random correlated standard normal variable is greater than $Z_\sigma$, then counterparty does not default, if it is less than $Z_\sigma$, the counterparty defaults.

When the counterparty does not default, its value at the end of year 1 is calculated by discounting the coupon payment and principal at the forward interest rates of its original credit rating. When the counterparty defaults, its value equals its principal value times the recovery rate.

After calculating the portfolio loss and repeating the procedure 10,000 times, the distribution of portfolio loss only due to default risk is obtained.

**CreditRisk+**

The data input for CreditRisk+ includes the bond's exposure, mean default rate and default rate volatility. The exposure for each bond is calculated as the bond market value at the end of year 1 minus its recovery amount. Assume the recovery rate is $R$, the exposure is then $[(1 - R) \times \text{Bond Value at year 1}]$. The mean default rate is obtained by calculating the average number of defaults from the 10,000 simulations for each bond. The default rate volatility data that was used was taken from Carty & Lieberman, Moody's Investors Service Global Credit Research.

Putting all the twenty bonds into one sector, the CreditRisk+ approach captures all of the concentration risk within the portfolio and excludes the benefits of diversification. Figure 4 presents the probability density for the number of defaults among the twenty
bonds. There is one AAA rated counterparty and two AA rated counterparties in this portfolio. According to S&P the probability of default within one year for both AA and AAA credit ratings is zero. As a result, the maximum number of default for the 20 bonds is 17. Since we are only interested in the portfolio loss, we can exclude the three bonds with zero mean default rate from the portfolio.

The default probability follows a Negative Binomial Distribution with \( \alpha_k = 1.65749451 \), \( \beta_k = 0.3091413 \), and \( p_k = 0.23614051 \). The 17 counterparties are separated into 10 different bands. The unit L is chosen to be $170,150. The exposure and the expected loss of each bond in terms of unit L is presented in Table 3. Figure 5 presents the probability distribution of portfolio loss and Table 4 provides the percentile data of the loss distribution.

IV. Results

Economic Capital

Bank capital is a bank's cushion against possible losses emanating from its other assets. Capital is needed in the form of liquid and secure assets. When losses arising from its other assets exceed the total amount of capital, the bank defaults and the excess loss is passed along to the creditors. Therefore, it is important for the bank to hold enough capital such that the probability of default is minimized. The amount of capital should be greater than the sum of its expected loss and unexpected loss. The expected loss is one of the costs of transacting business that gives rise to credit risk. Sufficient earnings should be generated through adequate pricing and provisioning to absorb any expected loss. However, the actual loss that occurs in any one period could be significantly higher than
the expected loss; therefore, the difference between the actual loss and the expected loss, which is known as the unexpected loss should be covered by a cushion called “economic capital”. The output of the credit risk portfolio models, the distribution of the portfolio credit loss, can be used to determine the level of economic capital required to cover the risk of unexpected credit losses. The percentiles of the loss distribution level provide a measure of economic capital for a required level of confidence. In order to capture a significant proportion of the tail of the portfolio loss distribution, the 99th percentile unexpected loss level over a one-year time horizon is a suitable definition for credit risk economic capital. Economic capital is the 99th percentile loss level minus the expected loss.

Before getting into the empirical comparison of CreditMetrics and CreditRisk+, let me provide the results of using the CreditMetrics approach to calculate the distribution of credit loss both with consideration of credit rating changes and without (Table 4). Note that when a counterparty defaults, the default value is based on its principal value.

The portfolio loss is substantially higher when we take account of credit rating changes. The economic capital suggested by CreditMetrics with credit rating changes is $5,225,023, which is 17.9 percent, or $792,823 more than that without considering rating changes. The corporation can easily get into financial distress with lack of this cushion $792,823.

The expected loss when considering credit rating changes is $618,368. The expected loss without credit rating changes effect is $536,557. Thus among the expected loss $618,368, about 87 percent of the loss comes from counterparties' default, and only
13 percent is due to credit rating changes. Therefore, default events cause much more significant value changes than any of the credit rating changes.

Table 5 provides the expected portfolio loss and the economic capital generated from both CreditMetrics and CreditRisk+. For CreditMetrics approach, I eliminate the credit rating change effect to make both models comparable, and when a counterparty defaults, the default value is calculated based on its market value at the end of year 1. Since the mean default rate of each counterparty is derived from the simulation, the expected portfolio loss is the same for both models which amounts to $445,891. The economic capital suggested by CreditMetrics is $3,731,446, 16.2 percent higher than that suggested by CreditRisk+, $3,212,334.

V. Conclusion

I have compared two of the major credit-risk portfolio models currently in use, CreditMetrics and CreditRisk+, both from a methodological and an empirical point of view. To facilitate the methodological comparison, I developed a framework to link the different input data employed by these two models. In particular, I showed that if asset values of different counterparties are related to some common economic factors, factor analysis may be used to transform the asset correlation data employed by CreditMetrics into the standard deviation data employed by CreditRisk+. This framework highlights the similarities as well as the differences between the two models.

Working with an example provided by J.P. Morgan (1997), I then implemented both the CreditMetrics and CreditRisk+ approaches, using the correlation data from J.P. Morgan (1997) and the standard deviation data from Credit Suisse Financial Products.
CreditMetrics is examined both with and without the inclusion of credit rating change risk. The inclusion of such risk results in a recommendation that the bank hold 17.9 percent more economic capital with rating change risk than without. This indicates the importance of taking such risk into account.

Since CreditRisk+ does not take account of credit rating change risk, I compared the CreditRisk+ recommendation to that of CreditMetrics when credit rating changes are excluded. Even in that case, the economic capital suggested by CreditMetrics is 16.2 percent higher than that suggested by CreditRisk+. When credit rating change risk is included, the amount of economic capital recommended by CreditMetrics is fully 39.9 percent higher than CreditRisk+'s recommendation. This indicates that the two models can yield very different results, and that a bank's model choice can make a substantial difference to its ability to avoid financial distress in the face of extreme loss.
References

“CreditMetrics - Technical document. The Benchmark for Understanding Credit Risk.”


Journal of Finance, 35, 1073-1103 (December 1980)

Figure 1. Probability of Default
Figure 2. Joint Probability of Default of Positively Correlated Counterparties A and B
Figure 3. Probability of Credit Rating Migration of a BBB-Rating Firm
Figure 4. Probability of number of defaults.
Figure 5. Distribution of Portfolio Loss
Table 1. Example Portfolio (20 Counterparties)

<table>
<thead>
<tr>
<th>Counterparty</th>
<th>Credit rating</th>
<th>Principal amount</th>
<th>Maturity (years)</th>
<th>Market value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AAA</td>
<td>7,000,000</td>
<td>3</td>
<td>7,821,049</td>
</tr>
<tr>
<td>2</td>
<td>AA</td>
<td>1,000,000</td>
<td>4</td>
<td>1,177,268</td>
</tr>
<tr>
<td>3</td>
<td>A</td>
<td>1,000,000</td>
<td>3</td>
<td>1,120,831</td>
</tr>
<tr>
<td>4</td>
<td>BBB</td>
<td>1,000,000</td>
<td>4</td>
<td>1,189,432</td>
</tr>
<tr>
<td>5</td>
<td>BB</td>
<td>1,000,000</td>
<td>3</td>
<td>1,154,641</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>1,000,000</td>
<td>4</td>
<td>1,263,523</td>
</tr>
<tr>
<td>7</td>
<td>CCC</td>
<td>1,000,000</td>
<td>2</td>
<td>1,127,628</td>
</tr>
<tr>
<td>8</td>
<td>A</td>
<td>10,000,000</td>
<td>8</td>
<td>14,229,071</td>
</tr>
<tr>
<td>9</td>
<td>BB</td>
<td>5,000,000</td>
<td>2</td>
<td>5,386,603</td>
</tr>
<tr>
<td>10</td>
<td>A</td>
<td>3,000,000</td>
<td>2</td>
<td>3,181,246</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>1000,000</td>
<td>4</td>
<td>1,181,246</td>
</tr>
<tr>
<td>12</td>
<td>A</td>
<td>2,000,000</td>
<td>5</td>
<td>2,483,922</td>
</tr>
<tr>
<td>13</td>
<td>B</td>
<td>600,000</td>
<td>3</td>
<td>705,409</td>
</tr>
<tr>
<td>14</td>
<td>B</td>
<td>1,000,000</td>
<td>2</td>
<td>1,087,841</td>
</tr>
<tr>
<td>15</td>
<td>B</td>
<td>3,000,000</td>
<td>2</td>
<td>3,263,523</td>
</tr>
<tr>
<td>16</td>
<td>B</td>
<td>2,000,000</td>
<td>4</td>
<td>2,527,046</td>
</tr>
<tr>
<td>17</td>
<td>BBB</td>
<td>1,000,000</td>
<td>6</td>
<td>1,315,720</td>
</tr>
<tr>
<td>18</td>
<td>BBB</td>
<td>8,000,000</td>
<td>5</td>
<td>10,020,611</td>
</tr>
<tr>
<td>19</td>
<td>BBB</td>
<td>1,000,000</td>
<td>3</td>
<td>1,118,178</td>
</tr>
<tr>
<td>20</td>
<td>AA</td>
<td>5,000,000</td>
<td>5</td>
<td>6,181,784</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>1.0</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>2</td>
<td>0.45</td>
<td>1.0</td>
<td>0.45</td>
<td>0.45</td>
</tr>
<tr>
<td>3</td>
<td>0.45</td>
<td>0.45</td>
<td>1.0</td>
<td>0.45</td>
</tr>
<tr>
<td>4</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>6</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>7</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>8</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>9</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>10</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>11</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>12</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>13</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>14</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>16</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>17</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>18</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>19</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>5</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>6</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>7</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>8</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>9</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>10</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>12</td>
<td>0.45</td>
<td>1</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>13</td>
<td>0.45</td>
<td>0.45</td>
<td>1</td>
<td>0.45</td>
<td>0.45</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>14</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>1</td>
<td>0.45</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>15</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>0.45</td>
<td>1</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.1</td>
</tr>
<tr>
<td>16</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>1</td>
<td>0.55</td>
<td>0.55</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>17</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.55</td>
<td>1</td>
<td>0.55</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>18</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.55</td>
<td>0.55</td>
<td>1</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>19</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 3. Exposure and Expected Loss Expressed as Multiples of the unit L, $170,150.

<table>
<thead>
<tr>
<th>Exposure ($v_j^{(k)}$)</th>
<th>Expected Loss ($e_j^{(k)}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.1086</td>
</tr>
<tr>
<td>3</td>
<td>0.2136</td>
</tr>
<tr>
<td>4</td>
<td>1.0184</td>
</tr>
<tr>
<td>8</td>
<td>0.456</td>
</tr>
<tr>
<td>9</td>
<td>0.0072</td>
</tr>
<tr>
<td>10</td>
<td>0.582</td>
</tr>
<tr>
<td>14</td>
<td>0.0154</td>
</tr>
<tr>
<td>16</td>
<td>0.2096</td>
</tr>
<tr>
<td>30</td>
<td>0.045</td>
</tr>
<tr>
<td>45</td>
<td>0.027</td>
</tr>
</tbody>
</table>
Table 4. Expected Loss and Economic Capital

CreditMetrics With Credit Rating Change Risk and Without
(Default value is based on the principal value)

<table>
<thead>
<tr>
<th></th>
<th>With</th>
<th>Without</th>
<th>Percentage Change(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Loss</td>
<td>618,368</td>
<td>536,557</td>
<td>15.2</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1,283,311</td>
<td>1,126,814</td>
<td>13.9</td>
</tr>
<tr>
<td>0.1%-Percentile Loss</td>
<td>11,571,010</td>
<td>11,296,582</td>
<td>2.4</td>
</tr>
<tr>
<td>1%-Percentile Loss</td>
<td>5,843,391</td>
<td>4,968,757</td>
<td>17.6</td>
</tr>
<tr>
<td>2.5%-Percentile Loss</td>
<td>3,961,583</td>
<td>3,552,550</td>
<td>11.5</td>
</tr>
<tr>
<td>5%-Percentile Loss</td>
<td>3,126,377</td>
<td>2,688,951</td>
<td>16.3</td>
</tr>
<tr>
<td>10%-Percentile Loss</td>
<td>2,046,552</td>
<td>1,643,434</td>
<td>24.5</td>
</tr>
<tr>
<td>Economic Capital</td>
<td>5,225,023</td>
<td>4,432,200</td>
<td>17.9</td>
</tr>
</tbody>
</table>

CreditMetrics and CreditRisk+
(Default value is based on the market value at the end of year 1)

<table>
<thead>
<tr>
<th></th>
<th>CreditMetrics (With)</th>
<th>CreditMetrics (Without)</th>
<th>CreditRisk+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expected Loss</td>
<td>527,703</td>
<td>445,891</td>
<td>445,891</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1,099,642</td>
<td>936,933</td>
<td>849,775</td>
</tr>
<tr>
<td>0.1%-Percentile Loss</td>
<td>9,440,926</td>
<td>9,000,249</td>
<td>6,720,925</td>
</tr>
<tr>
<td>1%-Percentile Loss</td>
<td>5,022,606</td>
<td>4,177,338</td>
<td>3,658,225</td>
</tr>
<tr>
<td>2.5%-Percentile Loss</td>
<td>3,479,557</td>
<td>2,965,025</td>
<td>2,637,325</td>
</tr>
<tr>
<td>5%-Percentile Loss</td>
<td>2,684,123</td>
<td>2,311,935</td>
<td>2,041,800</td>
</tr>
<tr>
<td>10%-Percentile Loss</td>
<td>1,797,034</td>
<td>1,302,891</td>
<td>1,531,350</td>
</tr>
<tr>
<td>Economic Capital</td>
<td>4,494,903</td>
<td>3,731,446</td>
<td>3,212,334</td>
</tr>
</tbody>
</table>
CreditMetrics versus CreditRisk+:
A Comparison of Two Credit Risk Portfolio Models
When Probabilities of Rating Migrations Are Known
I. Introduction

The fundamental concern when managing credit risk is the risk of loss considered on a portfolio basis. Consider a portfolio of assets, any one of which may default in the future. If an asset does default, its value will decrease by an amount that represents a loss to the holder of the portfolio. If we multiply the probability of default by the amount of the loss given default for each asset in the portfolio, we obtain the expected loss. The expected loss is easily calculated and should be reflected in asset pricing. Since the expected loss is not uncertain, it does not itself represent risk. However, the actual loss will in general differ from the expected loss. Risk in this context is the uncertainty in the actual loss that will occur. The riskiness of the portfolio arises from the deviation of the actual loss from the expected loss. This deviation is called the unexpected loss. The unexpected loss is not reflected in pricing and the possibility of such a loss threatening the solvency of the portfolio holder requires that a certain amount of money called economic capital be put aside to insure against insolvency.

One of the purposes of credit risk portfolio models is to measure credit risk at a portfolio level and to provide the decision maker with a recommendation about how much economic capital to hold to avoid financial distress without unnecessarily sacrificing investment opportunities. In He (1999), I compare two of the most popular credit risk portfolio models using a consistent set of input parameters: J. P. Morgan's CreditMetrics model and Credit Suisse First Boston’s CreditRisk+ model. In that paper, I used the transition matrix and the asset correlation data provided by J. P. Morgan (1997) to examine their models. When the possibility of credit rating changes is ignored, for the
portfolio examined in that paper, the CreditMetrics approach indicates that 16.2 percent more economic capital is required than the CreditRisk+ model does.

The biggest problem when comparing these two models is that we do not know the true transition matrix. Nor do we know how it may change from year to year. The data provided by J. P. Morgan are estimates of the true values based on historical data. They are used because the true values are unknown. When such data are used to compare the two credit risk portfolio models, all that we can learn is how each model would have performed over the historical period on which the estimates are based. We cannot necessarily conclude that the better performing model will continue its superior performance in the future. This problem is compounded by the fact that the historical estimates yield just one estimated transition matrix rather than allowing for changes in the matrix that might occur owing to changing economic conditions over the business cycle.

To resolve the problem, I propose creating a hypothetical world in which we know the true probability of credit rating migrations in each year, for each counterparty. Starting with a large pool of counterparties with different credit ratings and a "true" transition matrix for each year, I derive the "true" distribution of the portfolio loss, which is used as a benchmark required for the comparison of the two models. I then implement both models in the same way that they would be implemented in the real world. By simulating the credit rating evolution for each counterparty, I develop an estimated transition matrix and derive the standard deviation of default rates for each credit rating as well as the default correlations. With these data input, I derive the distributions of portfolio loss implied by both models. The performance of the two models is evaluated
by comparing the different distributions against the benchmark true distribution. If one method consistently dominates the other, the recommendation of the dominant model should be followed by decision makers to determine how much economic capital they should hold, conditional on the forecast future transition matrix.

**II. Modeling the Hypothetical World**

In the hypothetical world I assume there are three different states of the economy, representing good, average and poor economic condition. Each state of the economy has a different rating transition matrix. In the poor economic state the probability of downgrade and default is increased. In the good economic state the probability of downgrade and default is decreased. In modeling the evolution of the economy, each year one of the matrices is selected according to the phase of the business cycle. When the economy is in expansion, the good-year transition matrix provides the probabilities of upgrading, downgrading and default. When the economy is in recession, the poor-year transition matrix plays that role.

The three transition matrices are generated based on the historical data to represent the stylized facts. Based on the all-corporate average one-year rating transition matrix 1980-1998 provided in Moody's Investors Service (1999), I create the average-year transition matrix first. The poor- and good-year transition matrices are then developed by altering the average-year matrix in the appropriate manner. The business cycle is assumed to be a Markov process that is driven by a state-transition-matrix that takes account of the clustering of similar types of years in the economy.
In this hypothetical world, I start the economy in some state and evolve the state of the economy using the Markov transition model over a time horizon of 100 years. Each year the "true" transition matrix based on the current state of the economy is applied to a portfolio of firms. This procedure produces the "true" historical data representing the probabilities of credit migrations. By choosing a sufficiently long period the results are not sensitive to the initial assumed conditions. Based on the simulated data, I derive the standard deviations of default rates for each credit category. This becomes the data input for the CreditRisk+ approach. I also derive the average rating transition matrix as the data input for the CreditMetrics approach.

The transition matrices generated in the simulation have an important advantage; namely they can be used to obtain the default correlation data. J. P. Morgan (1997) uses the asset correlations because such default correlation data are unavailable. In He (1999), in order to facilitate the comparison of CreditMetrics and CreditRisk+, I use factor analysis to transform the asset correlation data employed by CreditMetrics into the standard deviation data employed by CreditRisk+. In this paper, when implementing the CreditMetrics approach, I use the default correlations directly. Since both the default correlation data and the standard deviation of default rates data are derived from the same historical data set, both models are compared using a consistent set of input parameters.

**Transition Matrices**

I employ three transition matrices representing three different economic conditions: good, average and poor economic states. Table 1 to Table 3 show the transition probabilities for the 3 matrices. For the average-year transition matrix, based on the all-corporate average one-year rating transition matrix 1980-1998 provided in
Moody's Investors Service (1999), I choose the probabilities to satisfy the following criteria. First, the probability of the credit rating remaining unchanged is the highest among the probabilities of credit migrations for each credit category. The absolute values of such probabilities for the investment-grade counterparties are higher than those for the speculative-grade, varying from 86 percent for AAA to 63 percent for CCC. Second, there is no chance for the counterparties rated AAA and AA to default, implying zero probabilities of default for both credit ratings. Third, the default probability increases when the credit rating deteriorates, ranging from 0.05 percent for A to 26.5 percent for CCC. Fourth, the probability of up/downgrading by one credit rating away from the original rating is higher than the probability of up/downgrading by two credit ratings away. For example, the probability of BBB upgrading to A is 7.5 percent, which is higher than the probability of BBB upgrading to AA, 0.55 percent. Fifth, the probabilities of each credit rating's migrations, which include upgrading, rating remaining unchanged, downgrading and default, sum up to 1.

Table 1. Average-Year Transition Matrix

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>86</td>
<td>10</td>
<td>3</td>
<td>0.6</td>
<td>0.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>1.5</td>
<td>87</td>
<td>9.5</td>
<td>0.85</td>
<td>0.55</td>
<td>0.45</td>
<td>0.15</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0.08</td>
<td>3</td>
<td>88.5</td>
<td>7.29</td>
<td>0.75</td>
<td>0.23</td>
<td>0.1</td>
<td>0.05</td>
</tr>
<tr>
<td>BBB</td>
<td>0.06</td>
<td>0.55</td>
<td>7.5</td>
<td>82.68</td>
<td>6.5</td>
<td>1.75</td>
<td>0.8</td>
<td>0.16</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0.3</td>
<td>0.75</td>
<td>7.2</td>
<td>8</td>
<td>8.6</td>
<td>1.65</td>
<td>1.5</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.1</td>
<td>0.51</td>
<td>0.75</td>
<td>8.8</td>
<td>78.5</td>
<td>3.84</td>
<td>7.5</td>
</tr>
<tr>
<td>CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>3.5</td>
<td>6.5</td>
<td>63</td>
<td>26.5</td>
</tr>
</tbody>
</table>

To develop the poor-year transition matrix, I alter the average-year matrix according to the following criteria. First, the probability of default increases for each
credit rating except for AAA and AA. The default probabilities vary from 0.1 percent for A to 35.05 percent for CCC. The absolute value increase in the probability is higher for the speculative-grade than for the investment-grade. In the poor-year state, the probability of default for A increases by 0.05 percent, while it increases by 8.55 percent for CCC. Second, the probability of credit rating remaining unchanged decreases for each credit category. Such decrease in the probability is larger for the speculative-grade than for the investment-grade. Making such an adjustment is in line with the intuitive notion that speculative-grade credit status varies more than investment-grade. For example, the probability of AAA remaining at AAA decreases by 3.6 percent in the poor year, while the probability of CCC remaining at CCC decreases by 6 percent. Third, the probabilities of downgrading for each credit category increase; for example, the probability of credit rating BBB downgrading to B increases from 6.5 percent in the average year to 9.5 percent in the poor year. Fourth, the probabilities of upgrading for each credit category decrease; for example, the probability of credit rating BBB upgrading to A decreases from 7.5 percent in the average year to 5.19 percent in the poor year.

<table>
<thead>
<tr>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>82.4</td>
<td>12</td>
<td>4</td>
<td>0.9</td>
<td>0.65</td>
<td>0.03</td>
<td>0.01</td>
</tr>
<tr>
<td>AA</td>
<td>1</td>
<td>83</td>
<td>12.94</td>
<td>1.5</td>
<td>0.8</td>
<td>0.55</td>
<td>0.2</td>
</tr>
<tr>
<td>A</td>
<td>0.04</td>
<td>2</td>
<td>84</td>
<td>10.5</td>
<td>2.3</td>
<td>0.81</td>
<td>0.25</td>
</tr>
<tr>
<td>BBB</td>
<td>0.03</td>
<td>0.35</td>
<td>5.19</td>
<td>78</td>
<td>9.5</td>
<td>4.2</td>
<td>2.5</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0.2</td>
<td>0.35</td>
<td>6.8</td>
<td>74.4</td>
<td>11.5</td>
<td>4.25</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.05</td>
<td>0.4</td>
<td>0.7</td>
<td>7.5</td>
<td>73</td>
<td>7.85</td>
</tr>
<tr>
<td>CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>2.2</td>
<td>5.5</td>
<td>57</td>
</tr>
</tbody>
</table>
To develop the good-year transition matrix, I make analogous adjustments in the opposite direction to those made to generate the poor-year matrix. First, the probability of default decreases for each rating. The absolute value decrease in the probability is larger for speculative-grade than for investment-grade. Second, the probability of the credit rating remaining unchanged increases, with the absolute value increase being higher for speculative-grade than for investment-grade. Third, the probabilities of downgrading for each rating decrease. Fourth, the probabilities of upgrading for each rating increase.

<table>
<thead>
<tr>
<th>Table 3. Good-Year Transition Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each cell contains the probability of transiting from the rating at the left of the row to the rating at the top of the column in a 1-year period.</td>
</tr>
<tr>
<td>AAA</td>
</tr>
<tr>
<td>AAA</td>
</tr>
<tr>
<td>AA</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>BBB</td>
</tr>
<tr>
<td>BB</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>CCC</td>
</tr>
</tbody>
</table>

**Business Cycle**

The fact that there exist clusters in the economy contributes significantly to the variation of default events. By and large, when the economy is expanding, a good economic year usually follows. On the other hand, when the economic state is on a downward trend, a poor economic year often follows. To model such clusters, I use a Markov model with a state-transition-matrix, a matrix that summarizes the conditional probabilities of each economic state occurring in the subsequent year. Because there are three different economic states in the hypothetical world, the state transition matrix is three by three. The three numbers in the first row give respectively the probability that a
good economy remains in the good state, that the good economy slows down to the average state, and that the good economy evolves to the poor state in the next year. The second row contains the same probabilities for an average economy, while the third row contains the same probabilities for a poor economy. When implementing CreditMetrics and CreditRisk+, I employed the following two state-transition-matrices:

\[
\begin{pmatrix}
0.8 & 0.2 & 0 \\
0.1 & 0.8 & 0.1 \\
0.1 & 0.2 & 0.7 \\
\end{pmatrix}
\] (2.1)

and

\[
\begin{pmatrix}
0.5 & 0.4 & 0.1 \\
0.2 & 0.6 & 0.2 \\
0.05 & 0.35 & 0.5 \\
\end{pmatrix}
\] (2.2).

If the economy is driven by the first state-transition-matrix (2.1), the economy tends to remain at its initial state. Using this matrix it generally takes longer for the economy to evolve to a different state. As a result the economic cycle is quite long. Compare to (2.1), an economy driven by the matrix (2.2) tends to fluctuate more frequently from state to state. By comparing the distribution of portfolio loss under the two different types of state-transition-matrices, we are able to see the influence of the business cycle on the results.

Given the state-transition-matrix for the business cycle, the evolution of the economy and the transition matrix for the credit rating migrations is selected by generating a random number from a uniform distribution ranging between 0 and 1. Assume we start with an average economy in year 1 and the business cycle is represented by the state transition matrix (2.1). If the random number generated is smaller than 0.1,
the economy in year 2 improves, and the good-year transition matrix is selected; if the random number is between 0.1 and 0.9, the economy in year 2 stays at the average state and the average-year transition matrix continues to be employed; if the random number is greater than 0.9, the economy deteriorates and the poor-year transition matrix is selected.

By the same token, given the economic state at year 2, based on the probabilities of state changes provided in the state-transition-matrix, we generate another random number uniformly distributed between 0 and 1 to determine the transition matrix applied to year 3 and so on.

**Default Correlation**

Another important data input for the CreditMetrics approach are the asset correlations that are utilized to determine the joint default behavior. The state of the economy at the start of the year is used to select transition matrix for the credit rating migrations for that year. The transition matrix for the credit rating migrations is then used to determine the year-end credit rating status of each counter party. Once the year and rating is determined the year-end bond values for a single counterparty are calculated. This is done for every counterparty and results in the value of the portfolio loss at the end of the year. The distribution of the portfolio loss is obtained by simulating this year-end portfolio loss many times. In each simulation every counterparty’s rating evolution is simulated.

He (1999) explains how J. P. Morgan (1997) obtains the counterparty’s year-end credit rating status by using asset correlations instead of default correlations. However, for the two-state CreditMetrics approach that considers only default risk and ignores
credit rating change risk, there is a shortcut to obtain the year-end portfolio value by using the default correlations directly. For the two-state CreditMetrics model for which the states are “default” and “no default”, the year-end portfolio value can be modeled by a set of Binomial random variables. The Binomial variables take the value of 1 if default occurs and 0 if there is no default. When a counterparty's Binomial random variable equals 0, the asset's value is simply its market value at the end of the year. When the Binomial random variable equals 1, the asset's value is its market value times its recovery rate.

Ideally, the pairwise default correlations calculated from the hypothetical world would be used to generate the correlated Binomial random variables directly. However, as explained in Appendix 1, if there are N counterparties, not only is it necessary to solve $2^N$ polynomial equations, but it is also necessary to track down $2^N$ probabilities to generate the Binomial random variables. Thus it is not practical to proceed in this manner. Instead, I utilize the idea of Merton's model that default is triggered by the asset value falling below the liability value. Assume that the asset return follows a Normal distribution, when it falls below a threshold value, default occurs; otherwise, there is no default. Therefore, the Binomial random variables are driven by a set of correlated Normal random variables. This set of Normal random variables is related to the Binomial variables in such a way that when the Normal variables fall below a set of threshold values, the corresponding Binomial variables take the value of 1; otherwise the Binomial variables take the value of 0. In addition, the probabilities of any two Normal variables falling below their threshold values simultaneously is the same as the probabilities of the corresponding Binomial variables taking the value of 1 simultaneously.
Consider an example of two counterparties with probability of default \( p_i \) and \( p_j \), and default correlation, \( \rho_y^g \). Their year-end Binomial values \( D_i \) and \( D_j \), either 1 or 0, are derived from two Normal random variables \( Y_i \) and \( Y_j \) respectively. The threshold values for the Normal variables are \( c_i \) and \( c_j \) respectively, where \( c_i = N^{-1}(p_i) \), \( c_j = N^{-1}(p_j) \), in which \( N \) is the Standard Normal cumulative density function. There are two steps to generate the Binomial random variables \( D_i \) and \( D_j \). Step 1 is to generate two Standard Normal random variables \( Y_i \) and \( Y_j \) with correlation \( \rho_y^N \), such that

\[
M(c_i, c_j; \rho_y^N) = p_{ij} = \rho_y^g \sqrt{p_i(1-p_i)p_j(1-p_j)} + p_i p_j,
\]

where \( M(c_i, c_j; \rho_y^N) \) is the cumulative probability in a standardized bivariate normal distribution that the first variable is less than \( c_i \) and the second variable is less than \( c_j \) when the coefficient of correlation between the variables is \( \rho_y^N \) and \( p_{ij} \) is the probability that both counterparties \( i \) and \( j \) default. Thus, given the probabilities of default \( p_i \) and \( p_j \), and the default correlation \( \rho_y^g \), the correlation of the corresponding Standard Normal variable \( \rho_y^N \) can be derived. The second equality of the above equation comes directly from the properties of the Binomial distribution. Step 2 is to compare the correlated Standard Normal random variables generated from step 1 with their threshold values, which gives us the Binomial random variable we require, with mean \( p_i \) and \( p_j \), and correlation \( \rho_y^g \):

\[
D_i = \begin{cases} 
1 & \text{when } Y_i \leq c_i \\
0 & \text{when } Y_i > c_i
\end{cases}
\]

and

80
\[
D_j = \begin{cases} 
1 & \text{when } Y_j \leq c_j \\
0 & \text{when } Y_j > c_j
\end{cases}
\]

It is important to note that the correlation between counterparty defaults in a given year and the correlation of default rates are quite distinct concepts. The default correlations derived from the hypothetical world simulations are correlations of default rates. In contrast, the default correlations \( \rho_{ij}^D \) necessary to generate the Binomial random variables are correlations of counterparties defaults. This is also the default correlation commonly referred to in the literature. Generating Binomial random variables therefore requires that the correlations of default rates derived from the hypothetical world be transformed into the correlations of counterparties defaults. Appendix 2 demonstrates how this transformation is accomplished. The default correlation between any two credit ratings A and B is

\[
\bar{\rho}_{\text{AB}} = \frac{\bar{\rho}_{\text{AB}} \sigma_{\text{A}} \sigma_{\text{B}}}{\sqrt{\mu_{\text{A}}(1-\mu_{\text{A}})\mu_{\text{B}}(1-\mu_{\text{B}})}},
\]

(2.3)

where \( \bar{\rho}_{\text{AB}} \) is the correlation of default rates and \( \bar{\rho}_{\text{AB}} \) is the correlation of counterparties defaults.

J. P. Morgan (1997) shows the default correlation within a credit rating being calculated as follows:

\[
\bar{\rho}_{\text{within C}} = \frac{N_C \left( \frac{\sigma_{\text{C}}^2}{\mu_{\text{C}}(1-\mu_{\text{C}})} \right) - 1}{N_C - 1}.
\]

(2.4)

where \( \mu_{\text{C}} \) is the mean default rate for credit category C, \( \sigma_{\text{C}} \) is the standard deviation of default rate for credit category C, and \( N_C \) is the number of the counterparties rated as C. As shown in Appendix 2, equation (2.4) is just a special case of equation (2.3).
Since the historical data are derived from the "true" transition matrices, the data input required for both CreditMetrics and CreditRisk+ are "true". Thus the distribution of the portfolio loss derived by using the "true" data input is "true" as well. This portfolio loss distribution is used as the benchmark to gauge the performance of the two models.

### III Simulating the Real World

In reality we can only observe up to 30 years of historical data and the transition matrices are not known. What practitioners observe each year in the market are the counterparties' different credit ratings, and whether each counterparty has been upgraded or downgraded, has defaulted or simply remained at its original credit rating.

To simulate such a real world, I select 30-year time horizon and 5000 counterparties. I simulate the evolution of the business cycle in the same way I did in the hypothetical world. This is to ensure that the simulated real world is consistent with the hypothetical world I create. Consequently, the "true" distribution of portfolio loss I derive in the hypothetical world can be used as a benchmark requested for the comparison of the two models. Based on each year's transition matrix, the evolution of credit status for each counterparty at the end of the year is determined. Given this information, I then calculate the estimated probabilities of credit migrations for each category. These are fractions in which the numerators represent the numbers of counterparties that migrate to different categories within the year and the denominators represent the total number of counterparties in that category. For example, at the beginning of year $t$, we have 500 BBB-rated counterparties. At the end of year, as shown in the second row of Table 4, 420
out of 500 remain at BBB, 37 are upgraded to A, 3 are upgraded to AA, 30 are
downgraded to BB, 8 are downgraded to B, and 2 counterparties default.

<table>
<thead>
<tr>
<th>Number of BBB firms</th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>Default</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>3</td>
<td>37</td>
<td>420</td>
<td>30</td>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Probabilities of Migrations</td>
<td>0.00</td>
<td>0.60</td>
<td>7.40</td>
<td>84.00</td>
<td>6.00</td>
<td>1.60</td>
<td>0.00</td>
<td>0.40</td>
</tr>
</tbody>
</table>

The estimated probabilities of migrations are calculated in the third row of table 4. For example, there are 37 BBB counterparties upgraded to A. Dividing 37 by the total of 500 BBB counterparties yields the probability of BBB upgrading to A as 7.4 percent. Since there are 2 counterparties default, the probability of default is 0.4 percent (or 2/500).

After repeating the above simulation over the time horizon of 30 years, I obtain the simulated historical data. I then derive the average transition matrix, the default correlations and the standard deviation of default rates as the data input for CreditMetrics and CreditRisk+. After implementing these two credit risk portfolio models, I obtain two different distributions of the portfolio loss. I then compare the two different results with the benchmark derived in the hypothetical world to examine the performance of the two models. If one model consistently dominates the other, the dominant model should be employed by banks to determine their economic capital. On the other hand, if neither model consistently dominates, it is impossible to make any such recommendation.

In the simulated real world, the changes in economic conditions together with the resulting changes in the transition matrices over the business cycle produce one
dimension of uncertainty affecting the default rates. The following two subsections provide a detailed illustration of two other sources of randomness, namely the rounding error resulting from replacement of defaulting firms and the randomness related directly to the transition matrices themselves. After simulating both this real world and the hypothetical world discussed above, I am able to decompose the uncertainty related to the default rates into its three source components.

**Rounding Error Resulting From Replacement**

Each year, there are some counterparties disappear from the portfolio due to default and some new counterparties are infused into the portfolio. I begin with a pool of 5000 counterparties which is close to the number of issuers that held Moody’s ratings on January 1 1999 as reported in Moody’s Investors Service (1999). In order to mimic the recent historical credit quality trends reported in the same document, I choose the number of counterparties in each credit category at year 1 so that the proportion of speculative-grade ratings in the portfolio is 40 percent, while the proportion of A- and BBB-rated counterparties is 45 percent, and that for AAA- and AA-rated is 15 percent. Table 5 shows the number of counterparties in each credit category.
Each year there are some counterparties that default. The same number of new counterparties is put into the portfolio at the beginning of the following year to keep the total number of counterparties in the portfolio unchanged. The credit rating of the new counterparties are determined by a set of proportions that is indicated in the second column of Table 6. For example, 20 percent of the new counterparties added into the portfolio each year are rated BBB, and 65 percent of them are rated B. Strikingly, the proportions of the new firms added into the portfolio for credit rating A and BB are zero. This is due to a special property of the transition matrices. The probabilities of remaining at their original credit rating for the counterparties from A or BB and the probabilities of upgrading or downgrading to A or BB for the counterparties from other ratings are especially high. In order to keep the number of the counterparties from each credit category relatively stable over years, it is necessary to keep the number of new
counterparties rated either A or BB extremely low.‡ A 30-year simulation of the credit quality trend illustrates that the proportion data provided in Table 5 is consistent with the historical trend provided in Moody's Investor's Service (1999).

Given the proportions of new counterparties with different credit ratings added to the portfolio, I multiply each proportion by the total number of defaults occurred the previous year. This yields the number of new counterparties in each credit category that join the portfolio in the current year. However, due to rounding error, the total number of new counterparties may not be exactly the same as the number of defaults. Consider an example where 955 counterparties default in one year. Then, there should be 955 new counterparties added into the portfolio. From the second column in Table 2 we know that 20 percent of the new counterparties are rated BBB. Thus 191 (955 × 20%) of the new counterparties are BBB rating. Notice however 5 percent of the new firms are rated AAA, which makes the number of new AAA counterparties 47.75 (955 × 5%), which is not an integer. Thus this number is rounded up to 48 new counterparties from AAA. Due to the rounding approximation, the number of new counterparties that replace the default counterparties might not be exactly the same as the number that default. In the above example, 955 counterparties default, while 956 new counterparties were added to the portfolio. The third column in Table 6 provides the rounded number of the counterparties from each credit rating.

‡ The interested readers may want to try adding up each column of a transition matrix to get a quantitative sense of this special property of the transition matrix.
Table 6

Percentage of new counterparties put into the portfolio from each credit rating. The final column shows the number of new additions that are created if 955 firms default.

<table>
<thead>
<tr>
<th>Credit Rating</th>
<th>Proportion</th>
<th>Out of 955 new firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.05</td>
<td>48</td>
</tr>
<tr>
<td>AA</td>
<td>0.05</td>
<td>48</td>
</tr>
<tr>
<td>A</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>BBB</td>
<td>0.20</td>
<td>191</td>
</tr>
<tr>
<td>BB</td>
<td>0.00</td>
<td>0</td>
</tr>
<tr>
<td>B</td>
<td>0.65</td>
<td>621</td>
</tr>
<tr>
<td>CCC</td>
<td>0.05</td>
<td>48</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1.00</strong></td>
<td><strong>956</strong></td>
</tr>
</tbody>
</table>

Randomness Related to Transition Matrix

At the beginning of each year there are approximately 5000 counterparties. At the end of each year, the credit rating of each counterparty is determined by the transition matrix for that year and the random numbers drawn to determine the evolution of the ratings. In the hypothetical world, we know the exact transition matrix applied to each year. In the real world, the transition matrix is not known. The information that is obtainable in the financial market is the credit status of each counterparty at the beginning of the year, and the evolved credit status of each counterparty at the end of the year. The transition matrix is derived by observing the actual number of firms whose ratings change in a year. For each transition, the probability is set equal to the number of firms that make this transition divided by the total number of firms in the initial rating class. In order to generate a real world that is consistent with the hypothetical world, I simulate each counterparty's credit evolution in two steps. First of all, I draw the "true"
transition matrix based on the business cycle the same way as I did in the hypothetical world. Second, for each counterparty, I generate a uniformly distributed random variable between 0 and 1, and use it in conjunction with the transition matrix to determine its credit evolution.

This is best illustrated by the following example. Assume there are 4999 counterparties at the beginning of year $n$. The number of counterparties from each credit rating is indicated in the Table 7.

<table>
<thead>
<tr>
<th>Credit Rating</th>
<th>Number of Firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>212</td>
</tr>
<tr>
<td>AA</td>
<td>482</td>
</tr>
<tr>
<td>A</td>
<td>1146</td>
</tr>
<tr>
<td>BBB</td>
<td>990</td>
</tr>
<tr>
<td>BB</td>
<td>759</td>
</tr>
<tr>
<td>B</td>
<td>1095</td>
</tr>
<tr>
<td>CCC</td>
<td>315</td>
</tr>
<tr>
<td>Total</td>
<td>4999</td>
</tr>
</tbody>
</table>

Take the example of credit rating BBB. As shown, there are 990 BBB counterparties in the portfolio. Assume the economy at year $n$ is in expansion so that the good-year transition matrix is selected for the credit rating migration probabilities. The probabilities of BBB's credit rating migrations, the data from the fourth row of the transition matrix, are presented in the following Table 8.
The actual number of BBB-rated counterparties migration to another credit category is determined by generating 990 uniformly distributed random numbers between 0 and 1. If the random number is smaller than 0.08 percent, then the counterparty is upgraded to AAA; if the random number is between 0.08 and 0.73 percent, it is upgraded to AA; if the random number is between 8.43 and 97.61 percent, it remains at BBB and so on. If the random number is greater than 99.9 percent, then a default occurs. Table 9 indicates how the credit rating migrations for credit rating BBB counterparties are determined.

<table>
<thead>
<tr>
<th>Random Number X</th>
<th>Credit Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>X&lt;0.08 percent</td>
<td>Upgrading to AAA</td>
</tr>
<tr>
<td>0.08 percent &lt; X &lt; 0.73 percent</td>
<td>Upgrading to AA</td>
</tr>
<tr>
<td>0.73 percent &lt; X &lt; 8.43 percent</td>
<td>Upgrading to A</td>
</tr>
<tr>
<td>8.43 percent &lt; X &lt; 97.61 percent</td>
<td>Remaining at BBB</td>
</tr>
<tr>
<td>97.61 percent &lt; X &lt; 98.85 percent</td>
<td>Downgrading to BB</td>
</tr>
<tr>
<td>98.85 percent &lt; X &lt; 99.6 percent</td>
<td>Downgrading to B</td>
</tr>
<tr>
<td>99.6 percent &lt; X &lt; 99.9 percent</td>
<td>Downgrading to CCC</td>
</tr>
<tr>
<td>X &gt; 99.9 percent</td>
<td>Default</td>
</tr>
</tbody>
</table>

Dividing the number of counterparties in each credit category by the total number of BBB counterparties at the beginning of the year yields the actual percentage of BBB counterparties that are upgraded, downgraded and default. Due to the uncertainty of the random numbers generated between 0 and 1, these percentages will in general be slightly
different from the ex-ante probabilities in Table 8. Proceeding in a similar vein, we will obtain the actual percentage of credit rating migrations for other credit ratings.

We have now simulated a history of the real world. We can see the different sources of the uncertainty that produces variations in default rates. The standard deviation of default rates partly results from the transition matrix selection process that occurs each year representing random changes in the phase of the business cycle; it partly results from the variation of the ex-post probabilities of migrations presented in the transition matrix representing the uncertainty of the asset values; and the rest arises from the rounding error of the new counterparties which replace the default counterparties each year.

**IV. Results**

I use the same portfolio as the one in He (1999) that contains 20 counterparties to compare the CreditMetrics and CreditRisk+ approaches. In the hypothetical world, the results of the “true” distributions of the portfolio loss calculated by each of the two models are presented in Table 10. The CreditMetrics and CreditRisk+ results are rather close to each other. The economic capital recommended by both approaches is approximately 2.8 million. However, CreditRisk+ has a lower standard deviation of portfolio loss.
This result is obtained by employing the state-transition-matrix (2.1). This generates an economy that tends to remain at its initial state. I also implement these two models using different state-transition-matrices to investigate the influence of the business cycle on the results. When employing the state-transition matrix (2.2) that causes the economy to fluctuate more frequently from state to state, the required economic capital is about 3.0 million. This result is quite intuitive; when the economy is more volatile, more capital should be put aside to guard against default risk.

To examine the performance of the two models, the portfolio loss distributions implied by both models are derived through simulating particular "real-world" outcomes and then compared against the benchmark, the true distribution derived in the hypothetical world. This examination is repeated 30 times. In general, the CreditRisk+ model recommends that more economic capital be held. One third of the time, both models produce results that are not significantly different from the benchmark, indicating that both models perform equally well. Slightly less than one half of the time the CreditMetrics model performs better in the sense that it produces an economic capital requirement closer to the benchmark requirement. One fifth of the time, the CreditRisk+
model performs better, producing an economic capital requirement closer to the benchmark.

From a statistical point of view, the CreditMetrics approach generally performs better. The statistical mean of the economic capital suggested by CreditMetrics in the real world simulations is 2.789 million, which is 1.48 percent away from the benchmark. In contrast, the statistical mean of the CreditRisk+ real world simulations is 2.917 million, which is 2.17 percent away. However, CreditRisk+ tends to overestimate and CreditMetrics to underestimate the economic capital requirement. Therefore, employing the CreditMetrics recommendation may result in greater damage when extreme loss occurs in the sense that there may not be enough capital to cover the loss. However, employing CreditRisk+’s approach may produce too much caution, causing more capital than necessary to be kept aside, preventing banks from taking advantages of beneficial investment opportunities.

V. Conclusion

I have compared two of the major credit risk portfolio models used by two prominent financial companies: J. P. Morgan’s CreditMetrics and Credit Swiss First Boston’s CreditRisk+, assuming the probabilities of credit rating migrations are known. By constructing a hypothetical world and deriving the “true” distribution of the portfolio loss as a benchmark, I then compare the two models against the benchmark to evaluate the models’ performance.
When implementing the CreditMetrics approach, I used the default correlation data directly to incorporate the joint default behavior. Given the known transition matrices and the implied counterparty migration probabilities, the standard deviation of default rates and the default correlation data can be derived directly from the same set of historical data. This has two advantages. First of all, since credit-rating-change-risk is omitted from the CreditMetrics approach, the default correlation data are sufficient for modeling the joint default behavior. Secondly, CreditMetrics and CreditRisk+ are easily comparable because the data used to represent the joint default behavior of the two models are consistent. This consistency eliminates the risk of modeling errors when transforming the asset correlation data into the standard deviation of default rates for the purpose of comparison. It also eliminates the necessity of finding an appropriate proxy to derive the asset correlations.

The results indicate that neither of the two models is consistently superior. Despite the fact shown in He (1999) that the two models yield different results, neither dominates the other. Therefore, banks need to consider carefully when making a choice between the two. Which model to employ depends on factors such as general economic conditions, the cost of implementing different models, the alternative investment opportunities available and each bank's own preferences. In general, forecasters who combine their intuition with the recommendations of models are better forecasters than those who blindly follow the models are. This is certainly true in the case of credit risk management.
References


Appendix 1. Generating Correlated Binomial Variables

First of all, consider an example of two Binomial variables \( D_A \) and \( D_B \) with correlation \( \rho_{AB} \). The probability of \( D_A \) taking the value 1 is \( P_A \) and the probability of \( D_B \) taking the value 1 is \( P_B \). Table A1 indicates the joint probabilities of the two variables.

<table>
<thead>
<tr>
<th>( D_A )</th>
<th>( D_B = 1 )</th>
<th>( D_B = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( P_{11} )</td>
<td>( P_{10} )</td>
</tr>
<tr>
<td>0</td>
<td>( P_{01} )</td>
<td>( P_{00} )</td>
</tr>
</tbody>
</table>

A set of polynomial equations can be set up as follows:

\[
P_A = P_{11} + P_{10}
\]

\[
P_B = P_{11} + P_{01}
\]

\[
P_{11} = \rho_{AB} \sqrt{P_A (1-P_A) P_B (1-P_B)} + P_A P_B
\]

\[
P_{11} + P_{10} + P_{01} + P_{00} = 1
\]

There are four equations and the four joint probabilities are unknown, thus the joint probabilities can be obtained by solving the above polynomial equations. Given \( P_{11}, P_{10}, P_{01} \) and \( P_{00} \) solved, we generate \( D_A \) first, then conditional on \( D_A \) we generate \( D_B \). To obtain \( D_A \), we generate a uniformly distributed random variable between 0 and 1. If the random variable is smaller than \( P_A \), \( D_A \) equals 1; otherwise, \( D_A \) equals 0. To obtain \( D_B \), we generate a second uniformly distributed random variable. Conditional on \( D_A \) being 1, we compare this random variable with \( \frac{P_{11}}{P_{11} + P_{10}} \), if it is smaller than \( \frac{P_{11}}{P_{11} + P_{10}} \), \( D_B \) equals 1; otherwise \( D_B \) equals 0. Conditional on \( D_A \) being 0, we compare the generated random variable with \( \frac{P_{01}}{P_{01} + P_{00}} \), if it is smaller than \( \frac{P_{01}}{P_{01} + P_{00}} \), \( D_B \) equals 1; otherwise \( D_B \) equals 0.
Now extend the case to three Binomial random variables $D_A, D_B$ and $D_C$ with the pairwise correlations $\rho_{AB}, \rho_{AC}$ and $\rho_{BC}$. The probabilities of taking the value 1 for three of them are $P_A, P_B$ and $P_C$. The polynomial equations are

$$P_A = P_{111} + P_{110} + P_{101} + P_{000} \quad (A1.b1)$$

$$P_B = P_{111} + P_{110} + P_{011} + P_{010} \quad (A1.b2)$$

$$P_C = P_{111} + P_{101} + P_{011} + P_{001} \quad (A1.b3)$$

$$P_{111} + P_{110} = \rho_{AB} \sqrt{P_A (1-P_A) P_B (1-P_B)} + P_A P_B \quad (A1.b4)$$

$$P_{111} + P_{101} = \rho_{AC} \sqrt{P_A (1-P_A) P_C (1-P_C)} + P_A P_C \quad (A1.b5)$$

$$P_{111} + P_{011} = \rho_{BC} \sqrt{P_B (1-P_B) P_C (1-P_C)} + P_B P_C \quad (A1.b6)$$

$$P_{111} = \rho_{AB} \sqrt{P_A (1-P_A) P_B (1-P_B)} + \rho_{BC} \sqrt{P_B (1-P_B) P_C (1-P_C)} + P_A P_B \quad (A1.b7)$$

$$P_{111} + P_{110} + P_{101} + P_{010} + P_{011} + P_{001} + P_{000} = 1 \quad (A1.b8)$$

where $P_{ijk}$ is the probability of $A$ taking the value $i$, $B$ taking the value $j$ and $C$ taking the value $k$. For example, $P_{101}$ is the probability of $A$ being 1, $B$ being 0 and $C$ being 1. In equation (A1.b7), we define the defaults $A$ and $B$ as a single event $AB$. Thus, the joint probability of three events $A$, $B$ and $C$ defaulting can be considered as the joint probability of two events defaulting, $AB$ and $C$. We can similarly combine defaults $A$ and $C$ as a single event and calculate the joint probability of this combined event with event $B$. No matter which pair of counterparties to combine, we find the same joint probability $P_{111}$.

There are eight unknowns and eight equations for the three counterparties case. By solving the polynomial equations (A1.b1) to (A1.b8), we can obtain the probabilities
\( P_{111}, P_{110}, P_{101}, P_{100}, P_{011}, P_{010}, P_{001}, P_{000} \). To generate these three correlated Binomial variables, we generate \( D_A \) first. Based on \( D_A \), we then generate \( D_B \). Based on both \( D_A \) and \( D_B \), we finally generate \( D_C \). Obtaining \( D_A \) is the same as the case when generating two correlated Binomial variables. To obtain \( D_B \), we generate a uniformly distributed random variable, say \( x_2 \). Conditional on \( D_A \) being 1, we compare \( x_2 \) with \( \frac{P_{11}}{P_{11} + P_{10}} \) where

\[
P_{11} = P_{111} + P_{110} \quad \text{and} \quad P_{10} = P_{101} + P_{100}.
\]

If \( x_2 \) is smaller than \( \frac{P_{11}}{P_{11} + P_{10}} \), \( D_B \) equals 1; otherwise \( D_B \) equals 0. Conditional on \( D_A \) being 0, we compare \( x_2 \) with \( \frac{P_{01}}{P_{01} + P_{00}} \) where

\[
P_{01} = P_{011} + P_{010} \quad \text{and} \quad P_{00} = P_{001} + P_{000}.
\]

If \( x_2 \) is smaller than \( \frac{P_{01}}{P_{01} + P_{00}} \), \( D_B \) equals 1; otherwise \( D_B \) equals 0. To generate \( D_C \), we generate another uniformly distributed random variable, \( x_3 \), then

Conditional on \( D_A = 1 \) and \( D_B = 1 \),

\[
D_C = \begin{cases} 
1 & \text{if } x_3 < \frac{P_{111}}{P_{111} + P_{110}} \\
0 & \text{otherwise}
\end{cases}
\]

Conditional on \( D_A = 1 \) and \( D_B = 0 \),

\[
D_C = \begin{cases} 
1 & \text{if } x_3 < \frac{P_{101}}{P_{101} + P_{100}} \\
0 & \text{otherwise}
\end{cases}
\]

Conditional on \( D_A = 0 \) and \( D_B = 1 \),

\[
D_C = \begin{cases} 
1 & \text{if } x_3 < \frac{P_{011}}{P_{011} + P_{010}} \\
0 & \text{otherwise}
\end{cases}
\]

97
Conditional on $D_A = 0$ and $D_B = 0$,

$$D_C = \begin{cases} 1 & \text{if } x_i < \frac{P_{001}}{P_{001} + P_{000}} \\ 0 & \text{otherwise} \end{cases}$$

In general, to generate $N$ Binomial variables, we need to solve $2^N$ equations to obtain $2^N$ probabilities. Therefore, given the pairwise correlations between Binomial random variables, it is not practical to generate the correlated Binomial random variables directly.
Appendix 2. Derivation of Default Correlations Across Credit Ratings

There are two groups of counterparties rated, say, A and B. The number of the counterparties that each group contain are $N_A$ and $N_B$ respectively. Assume the counterparties from the same credit category possess the identical default rate which are $\mu_A$ and $\mu_B$ respectively. Let $X_i$ be the random variable that is either 1 or 0 according to each counterparty’s default event realization from credit rating A and $X_j$ be that from credit rating B. The mean default rate $\mu_A$, $\mu_B$ and binomial default standard deviations $\sigma(X_i)$, $\sigma(X_j)$ are defined as follows,

$$X_i = \begin{cases} 
1 & \text{if counterparty } i \text{ defaults} \\
0 & \text{otherwise}
\end{cases}$$

and

$$X_j = \begin{cases} 
1 & \text{if counterparty } j \text{ defaults} \\
0 & \text{otherwise}
\end{cases}$$

$$\mu_A = \frac{1}{N_A} \sum_i X_i \quad \text{and} \quad \mu_B = \frac{1}{N_B} \sum_j X_j$$

$$\sigma(X_i) = \sqrt{\mu_A(1-\mu_A)} \quad \text{and} \quad \sigma(X_j) = \sqrt{\mu_B(1-\mu_B)}.$$

Let $D_A$ and $D_B$ represent the number of defaults from the two credit categories, where

$$D_A = \sum_i X_i \quad \text{and} \quad D_B = \sum_j X_j,$$

and D represents the total number of defaults form both categories, $D = D_A + D_B$. Then the variance of D is:
\[
Var(D) = Var\left( \sum_{k}^{N_A} X_k \right) = \sum_{i}^{N_A} \sum_{j}^{N_B} \rho_{ij} \sigma_i \sigma_j
\]

\[
= \left( N_A + \sum_{i}^{N_A} \sum_{i \neq j} \rho_{ij} \right) \sigma_i^2(\mu_i - \mu_A)^2 + \left( N_B + \sum_{i}^{N_B} \sum_{i \neq j} \rho_{ij} \right) \sigma_j^2(\mu_j - \mu_B)^2 + 2N_A N_B \bar{\rho}_{AB} \sigma_i \sigma_j
\]

\[
= \left[ N_A + (N_A^2 - N_A) \bar{\rho}_A \right] [\mu_A (1 - \mu_A)] + \left[ N_B + (N_B^2 - N_B) \bar{\rho}_B \right] [\mu_B (1 - \mu_B)] + 2N_A N_B \bar{\rho}_{AB} \sigma_i \sigma_j
\]

(A2.1)

where \( \bar{\rho}_{AB} \) is the average default correlation between credit rating A and B; while \( \bar{\rho}_A \) and \( \bar{\rho}_B \) are the average default correlation within credit categories A and B respectively which are defined in J. P. Morgan (1997) as:

\[
\bar{\rho}_{\text{within}} = \left[ 1 - \sum_{i}^{N_A} \sum_{j \neq i} \rho_{ij} \right] \frac{1}{(N^2 - N)}
\]

where \( \rho_{ij} \) is the correlation of any two counterparties within a credit category.

Since the variance of the total number of defaults can also be derived as follows,

\[
Var(\sum_{i}^{N_A} X_i) = Var(\sum_{i}^{N_A} X_i) + Var(\sum_{i}^{N_B} X_i) + 2 \text{cov}(\sum_{i}^{N_A} X_i, \sum_{i}^{N_B} X_i)
\]

\[
= N_A^2 \text{Var} \left( \frac{\sum_{i}^{N_A} X_i}{N_A} \right) + N_B^2 \text{Var} \left( \frac{\sum_{i}^{N_B} X_i}{N_B} \right) + 2N_A N_B \bar{\rho}_{AB} \sigma \left( \frac{\sum_{i}^{N_A} X_i}{N_A} \sigma \left( \frac{\sum_{i}^{N_B} X_i}{N_B} \right)
\]

(A2.2)

where \( \bar{\rho}_{AB} = \rho \left( \frac{\sum_{i}^{N_A} X_i}{N_A}, \frac{\sum_{i}^{N_B} X_i}{N_B} \right) \) which is the correlation of default rates across credit ratings A and B; \( \sigma_{\text{CrA}} \) and \( \sigma_{\text{CrB}} \) are the standard deviations of default rates for credit categories A and B respectively. Thus, the average default correlation across credit ratings can be derived by combining (A2.1) and (A2.2), that is
When we see $N_i = N_j = N$, $\mu_i = \mu_j = \mu_C$, $\bar{\rho}_A = \bar{\rho}_B$, and $\sigma^2_{C_{CC}} = \frac{\sigma^2_{C_{CL}} + \sigma^2_{C_{CB}}}{2}$,

Equation (A2.3) boils down to the average default correlation within a credit category which is derived in J. P. Morgan (1997),

\[
\bar{\rho}_{\text{within } C} = \frac{N_C \left( \frac{\sigma^2_{C_{CC}}}{\mu_C(1-\mu_C)} \right) - 1}{N_C - 1}, \tag{A2.4}
\]

where $\sigma^2_{C_{CC}}$ is the standard deviation of default rates for credit category $C$, and $N_C$ is the number of counterparties rated as $C$. Substituting equation (A2.4) into (A2.3), the average default correlation across credit categories is simplified as follows:

\[
\bar{\rho}_{AB} = \frac{\bar{\rho}_{AB} \sigma^2_{C_{CL}} \sigma^2_{C_{CB}}}{\sqrt{\mu_A(1-\mu_A)\mu_B(1-\mu_B)}} \tag{A2.5}.
\]

(A2.5) also indicates the relation between the correlation of counterparties defaults across credit ratings and the correlation of default rates across credit ratings.