



March 2000

Rating Methodology

Contact	Phone
New York Jorge R. Sobehart Roger M. Stein	1.212.553.1653

RATING METHODOLOGY

Moody's Public Firm Risk Model: A Hybrid Approach To Modeling Short Term Default Risk

*Jorge R. Sobehart, Roger M. Stein
Victoriya Mikityanskaya, Li Li*

William Thompson, Susan Heckman

© Copyright 2000 by Moody's Investors Service, Inc., 99 Church Street, New York, New York 10007. All rights reserved. **ALL INFORMATION CONTAINED HEREIN IS COPYRIGHTED IN THE NAME OF MOODY'S INVESTORS SERVICE, INC. ("MOODY'S"), AND NONE OF SUCH INFORMATION MAY BE COPIED OR OTHERWISE REPRODUCED, REPACKAGED, FURTHER TRANSMITTED, TRANSFERRED, DISSEMINATED, REDISTRIBUTED OR RESOLD, OR STORED FOR SUBSEQUENT USE FOR ANY SUCH PURPOSE, IN WHOLE OR IN PART, IN ANY FORM OR MANNER OR BY ANY MEANS WHATSOEVER, BY ANY PERSON WITHOUT MOODY'S PRIOR WRITTEN CONSENT.** All information contained herein is obtained by **MOODY'S** from sources believed by it to be accurate and reliable. Because of the possibility of human or mechanical error as well as other factors, however, such information is provided "as is" without warranty of any kind and **MOODY'S**, in particular, makes no representation or warranty, express or implied, as to the accuracy, timeliness, completeness, merchantability or fitness for any particular purpose of any such information. Under no circumstances shall **MOODY'S** have any liability to any person or entity for (a) any loss or damage in whole or in part caused by, resulting from, or relating to, any error (negligent or otherwise) or other circumstance or contingency within or outside the control of **MOODY'S** or any of its directors, officers, employees or agents in connection with the procurement, collection, compilation, analysis, interpretation, communication, publication or delivery of any such information, or (b) any direct, indirect, special, consequential, compensatory or incidental damages whatsoever (including without limitation, lost profits), even if **MOODY'S** is advised in advance of the possibility of such damages, resulting from the use of or inability to use, any such information. The credit ratings, if any, constituting part of the information contained herein are, and must be construed solely as, statements of opinion and not statements of fact or recommendations to purchase, sell or hold any securities. **NO WARRANTY, EXPRESS OR IMPLIED, AS TO THE ACCURACY, TIMELINESS, COMPLETENESS, MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE OF ANY SUCH RATING OR OTHER OPINION OR INFORMATION IS GIVEN OR MADE BY MOODY'S IN ANY FORM OR MANNER WHATSOEVER.** Each rating or other opinion must be weighed solely as one factor in any investment decision made by or on behalf of any user of the information contained herein, and each such user must accordingly make its own study and evaluation of each security and of each issuer and guarantor of, and each provider of credit support for, each security that it may consider purchasing, holding or selling. Pursuant to Section 17(b) of the Securities Act of 1933, **MOODY'S** hereby discloses that most issuers of debt securities (including corporate and municipal bonds, debentures, notes and commercial paper) and preferred stock rated by **MOODY'S** have, prior to assignment of any rating, agreed to pay to **MOODY'S** for appraisal and rating services rendered by it fees ranging from \$1,000 to \$1,500,000.

PRINTED IN U.S.A.

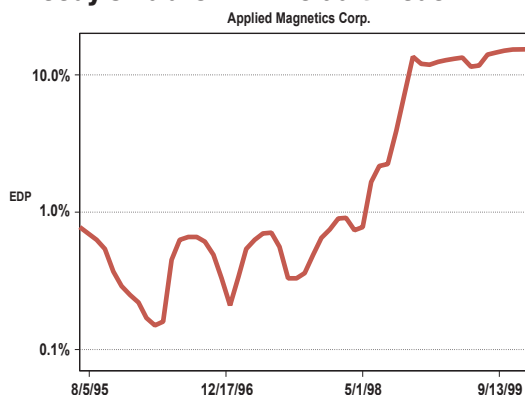
Summary	5
1. Introduction	6
2. Model Description	7
2.1 The Core Of Moody's Public Firm Model: A Contingent Claims View Of Credit Risk	8
2.2 Ratings And Financial Statement Variables In Moody's Public Firm Model	9
2.3 A Non-linear Modeling Approach	12
3. Model Performance And Benchmarking Of Moody's Model Against Other Models ..	13
3.1 Measuring Model Accuracy	14
3.1.1 Cumulative Accuracy Profiles (CAPs)	15
3.1.2 Accuracy Ratio (AR)	16
4. Conclusions	18
5. References	19
6. Appendices	21
6.1 Appendix A: Variable Selection And Model Construction	21
6.2 Appendix B: An Introduction To Model Performance And Validation	22
6.3 Appendix C: Estimated Default Probability	23
6.4 Appendix D: A Non-linear Approach To Modeling Default	25
6.5 Appendix E: Moody's Default Modeling Database	25

4

Moody's Rating Methodology

In this Rating Methodology piece, we report the results of research that Moody's Risk Management Services has done in modeling default risk for public firms. More precisely, we introduce Moody's Public Firm Risk Model and discuss its performance. We used Moody's proprietary data, financial statements and market information from a variety of commercial vendors to construct a default prediction model for public firms. The model incorporates ratings and financial statement data (as do statistical credit models) as well as market information in the form of a structural model (as do contingent claims models).

Exhibit 1. Moody's Public Firm Default Model: An Anecdotal Example



On January 7, 2000 Applied Magnetics Corporation, an unrated obligor, filed for protection under Chapter 11 of the US Bankruptcy Code in Santa Barbara. Applied Magnetics Corp. is a manufacturing firm specializing in magnetic recording heads. Several factors led to persistent and large losses for the firm including a protracted downturn in the drives industry due to excess capacity, pressure on average selling prices resulting from the migration to less expensive computers and a reduction in demand for both head stack assemblies and disk drive platters. Additionally, the late launch of certain products and significant capital expenditures resulted in operating difficulties. For reference, the EDP for Applied Magnetics Corp. crossed above the average EDP for the corporate universe around May 1998, about a year and a half before the actual default. The EDP increased steeply thereafter.

Moody's model provides a one-year estimated default probability (EDP) for individual issuers, such as Applied Magnetics Corp. above (**Exhibit 1**), using the following information as input:

- a variant of Merton's option theoretic model of firms;
- Moody's rating (when available);
- company financial statement information;
- additional equity market information; and
- macroeconomic variables which represent snapshots of the state of the economy and of specific industries.

The model can be used as an early warning system for monitoring short-term changes in the credit quality of corporate obligors. The current model is broad in scope; that is, it has the same structure across all industries. The model does, however, contain refinements to address particular features of individual industries. We are currently applying the model to U.S. non-financial firms.

In brief:

- We present a description of the current version of Moody's Public Firm model as well as an evaluation of its effectiveness at anticipating default events. More precisely, we describe our model, the models used as benchmarks, their performance, and the value added by Moody's model.
- Moody's Public Firm model performs better at anticipating defaults than a variety of linear, contingent claims and logistic regression models reported in the literature. We found that our model consistently outperforms other models when all models are benchmarked on the same comprehensive data set. This suggests that Moody's model can be used productively as an early warning system for credit risk assessment and portfolio risk management.

3. Our research has shown that Moody's hybrid model adds predictive value over a pure contingent claims approach derived from the Merton (1973, 1974) model.
4. We evaluated these alternative models using data from Moody's extensive proprietary default database and the credit model validation approach described in Sobehart, Keenan and Stein (2000). We performed comprehensive performance benchmarking across industries, time periods, and models. To our knowledge, it is the first time that such research has been undertaken using a comprehensive data set and a consistent methodology based on the information content of credit risk models.
5. In addition to looking at the number of defaults predicted, which we have found to be a limited measure of model performance, we focus our attention on two of many performance measures we used: Cumulative Accuracy Profiles (CAP) curves and Accuracy Ratios (AR).

1. Introduction

Credit risk can be defined as the potential that a borrower will fail to meet its obligations in accordance with agreed terms. For most individual and institutional investors, bonds and other tradable debt instruments are the main source of credit risk. In contrast, for banks, loans are the largest and most obvious source of credit risk. Banks need to manage the credit risk exposure in their entire portfolio as well as the risk in individual credits or transactions. Over the last years, a number of the world's largest financial institutions have developed advanced systems in an attempt to model their exposure to credit risk. Such models are intended to aid institutions in quantifying, monitoring and managing risk across their business lines.

Default models represent a strategic component of the set of quantitative tools required by financial institutions. A reliable default risk model for public firms performs a vital role by helping analysts to make informed credit decisions by associating default probability with borrower firms and counterparties. Such a model can be used as a monitoring tool for screening obligors, for performing risk/return analysis of credit portfolios or for capital allocation and loan pricing.

Moody's Public Firm model has been designed to act as an early warning system to monitor changes in the credit quality of corporate obligors. For each issuer, the model currently reports a one year *estimated default probability* (EDP, sometimes called an EDF) using as inputs ratings (when available), market data and financial statement information. The scope of the model is broad in that it is applicable across a wide variety of industries. Although it is structurally the same for all industries, the model accounts for the particular features of individual industries through adjustments to input data based on broad industry and macro-economic variables. With the exception of financial institutions, whose capital structures tend to be quite unique, we are currently applying the model to all industries and U.S. firms.

Moody's model is a hybrid one that combines two credit risk modeling approaches: (a) a structural model based on Merton's options-theoretic view of firms, and (b) a statistical model determined through empirical analysis of historical data.¹

The statistical approach, which is the most frequently found in the literature², maps a reduced set of financial variables and other information to a risk scale. The mapping acts as a statistical distillation of the historical data and can be used to discriminate between good and bad credits. The linear model introduced by Altman³, also known as Z-score, is an example of a reduced representation statistical model. The Altman model separates defaulting from non-defaulting firms based on the discriminatory power of a linear combination of financial ratios.

However, the analysis of historical financial statements may present an incomplete or distorted picture of the company's true financial condition. For a variety of reasons including the intrinsic conservatism of accounting principles, financial statements do not necessarily reflect the complete economic reality of the firm. Furthermore, accounting practices do not provide a means for expressing uncertainty about the

1. In practice, most models have attributes of both approaches. Most statistical models begin with some theoretical framework to aid in problem formulation, and most theoretical models rely on the use of statistics to determine the appropriate values for key parameters, such as volatility, or to map structural models to EDFs.

2. See, for example, Foster (1986).

3. See Altman (1968, 1993).

future since the fundamental principle is to “account” for all the items involved in the firm’s operations during every period precisely. Unfortunately, while financial statements provide information directly about a firm’s past, they are limiting in that they provide information only indirectly about its future.

The second modeling approach starts with a stylized mathematical representation of how the value of the firm evolves through time. The goal of this type of quantitative risk assessment is to represent the solvency of the issuer in the theoretical economic environment as accurately as possible, based on financial principles. The valuation of corporate liabilities using the contingent claims framework introduced by Merton⁴ is an example of a structural model. The usefulness of such an approach depends on how closely its assumptions and structure capture the true nature of the firm dynamics as well as the accuracy with which the model’s variables are estimated.

In particular, the Merton model relies heavily on economic theories about market efficiency. The model contains embedded assumptions about the comprehensiveness of the information contained in market data when used within the structure of the model. However, knowledge of the market information alone does not directly inform an investor as to a borrower’s creditworthiness.

For example, Merton’s original contingent claims model, and most subsequent refinements of it, does not contemplate cases in which firms default on their debt obligations due to severe liquidity problems. In addition, even in situations where market equity contains relatively complete information about a firm’s credit quality, this does not guarantee that a structural model will correctly capture and reflect that information⁵.

Thus, while market information can be extremely valuable, we have found it to be most useful when also coupled with fundamental information on the firm and its business environment. A detailed examination of a firm’s balance sheet, income statement and cash flows remains a critical component of any analytical risk assessment framework.

Moody’s approach reflects this perspective. Our model incorporates variants of both a contingent claims model and a statistical reduced form model using a non-linear regression approach. The key inputs to this hybrid model are:

- (a) agency rating when available,
- (b) modified version of the Merton model (expressed as a distance to default),
- (c) company financial statement information,
- (d) additional equity market information; and
- (e) macroeconomic variables that represent snapshots of the state of the economy or of specific industries which are used for preprocessing model inputs.

Ratings and default information are obtained from Moody’s proprietary rating and default databases for public firms. Fundamental financial and market information is collected from commercial vendors.

In this *Rating Methodology* we present Moody’s model, and rigorously examine the model’s ability to identify troubled firms. Section 2 presents the modeling approach and data used to build our Public Firm model. Our validation methodology and the model testing are discussed in Section 3. More precisely, we describe the models used as benchmarks, their performance, and the value added by Moody’s model. In section 4 we present our main conclusions.

2. Model Description

Moody’s Public Firm default model is a hybrid one that combines a contingent claims approach to credit risk with a statistical approach. The model we introduce here provides an estimated default probability (EDP) over a one-year horizon. The EDP (sometimes called an EDF) for individual firms is calculated based on financial statement data, capital markets information and Moody’s ratings (when available).

4. See Merton (1973, 1974).

5. For a technical discussion on what the market value of equity can and cannot tell analysts about the credit risk of corporate borrowers see Sobehart and Keenan (1999). Interested readers are also encouraged to refer to the papers cited at the end of this *Rating Methodology* for more detail on our approach.

Technically, our default risk model is a nonlinear regression model based on a combination of logistic regressions⁶ (sometimes referred to as an artificial neural network) and Bayesian inference. In the next several sections, we describe our approach and model building methodology.

2.1 THE CORE OF MOODY'S PUBLIC FIRM MODEL: A CONTINGENT CLAIMS VIEW OF CREDIT RISK

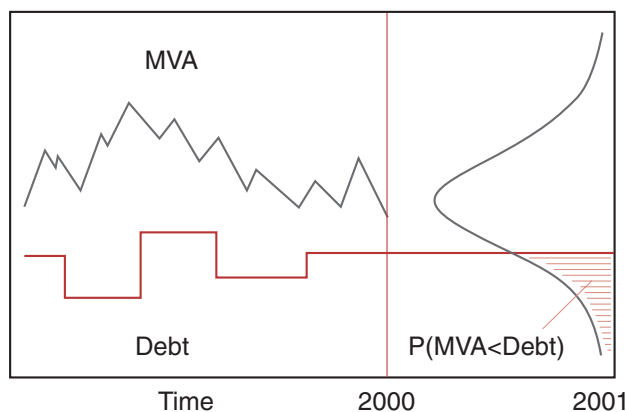
Central to Moody's model is a variant of Merton's (1973, 1974) analytical model of firm value. Conceptually, this model views a firm's equity as an option on the firm (held by shareholders) to either repay the debt obligations of the firm when they come due, or to abandon the firm without paying the obligations at that time. In this subsection, we present a brief overview of this approach.

In their seminal work on the valuation of options, Black and Scholes (1973) made the insightful observation that the common stock of a firm possesses option-like features with respect to the solvency of the firm. Merton ((1973) and (1974)) developed this idea even further deriving a contingent claims model based on the relation between the value of a firm's equity and the value of its debt. Fundamental to Merton's model is the idea that equity and debt could be considered as options on the value of the firm's assets.

To see this, consider the simple case of a firm that has a single debt obligation outstanding, such as a loan or a zero-coupon bond. The debt matures in one year, calling for a single payment of principal and interest at maturity. There are no interim interest payments during the year. Now consider the position of the stockholders. At the end of the year they have the choice of either repaying the debt-holders to retain control of the firm or, by not repaying the debt, going into default and surrendering ownership of the firm to the debt-holders.

Under this framework, the stockholders own an option to buy the firm from the debt-holders at an exercise price equal to the debt value, with an expiration date equal to the time to maturity of the debt. Since the stockholders' payoff is theoretically identical to that of an option, equity should be priced in the market as though it were an option, with an exercise price equal to the outstanding debt value and one year to expiration.

Exhibit 2. Simplified Schematic Of A Contingent Claims Model Of Credit Risk



Conceptually, the contingent claims model views the firm's equity as a call option with a strike price equal to the par value of the debt obligations. By estimating the probability that the market value of the assets (MVA) will be below the value of the debt (shaded region), this model can be used to estimate the probability of default. The market value of the firm's assets and its volatility are inferred from the firm's equity value.

In the Merton model the firm's susceptibility to financial distress is reflected by the *probability of the firm's assets being below the value of the promised payments* (Exhibit 2). In this particular situation, the firm would be unable to raise additional capital to cover its debts and would eventually default. Thus, the Merton model provides a simple theoretical framework for estimating the probability of default of obligors.

6. Moody's has been applying non-linear computational techniques to modeling various aspects of credit analysis for nearly a decade (e.g., Stein (1991) describes early work on other types of credit models).

Unfortunately, the original Merton model does not correlate well with observed market behavior⁷. As a result, the Merton model must be adapted to give realistic default probabilities. More precisely, the variant of the Merton model is not used to calculate default probabilities but rather to calculate the market value of the firm's assets and volatility from equity prices, which are related to the default probability.

The default probability is a function of the number of standard deviations that the value of the firm's assets must drop in order to reach the default point. This number is called the *Distance to Default*⁸. The concept of *Distance to Default* plays a key role in Moody's Public Firm Default model.

In pure market based models, the Distance to Default is usually mapped to historical default probabilities, a practice that transforms the structural model into a statistical one⁹.

2.2 RATINGS AND FINANCIAL STATEMENT VARIABLES IN MOODY'S PUBLIC FIRM MODEL

Although a contingent claims framework forms the core of our model, we have found it most useful when combined with additional financial data and ratings¹⁰. Therefore, Moody's default model is a hybrid that also incorporates financial statement information and ratings, when available. The final output of our model is based on the statistical relationships to default of financial statement information, market information, ratings (when they exist) and a variant of Merton's contingent claims model expressed as a distance to default.

When firms are rated, the rating has proven to be an extremely useful indicator of long term solvency.¹¹ Because the rating is a multidimensional aggregation of credit information, it provides valuable information directly on the credit quality of an issuer. Accordingly, it is another important input to our model. When a firm has no rating, the Public Firm model uses an estimated rating based on a proprietary rating model in its place.

Since the availability of reliable data is always a limiting factor in model building, selecting additional factors for the model (beyond rating and the distance to default) reduces the number of usable observations, as well as the applicability of the model to new firms. Moody's model was optimized with the additional constraint of imposing minimum data requirements on users. We do this to provide maximum coverage of the universe of unrated public firms.

Because of the non-linear nature of the model, and also because the credit histories for individual obligors are highly correlated in time, it is inappropriate to rely on many of the commonly used methods for variable selection.¹² Instead, we evaluated the relative significance of the variables using a variety of non-parametric methods based on resampling.¹³

To select the minimum set of explanatory variables, we began with a set of fundamental accounting and market-based variables in addition to our contingent claims model. Next, we used conventional statistical techniques such as linear and nonlinear discriminant analysis, tempered with input from Moody's industry analysts, to determine the most likely candidate variables (while still trying to maximize the coverage of public firms). Finally, we used resampling methods to conduct extensive sensitivity analysis. We did this by bootstrapping different performance metrics for various candidate variable sets.¹⁴

As mentioned above, the objective of the analysis was to maximize the model's accuracy while simultaneously keeping the data requirements to a minimum. The final set of selected financial and market variables are described in **Exhibit 3**.

7. If the Merton model were, in fact, a consistent representation of actual equity and bond price dynamics, it should result in credit spreads similar to those observed in the bond market. This is not, however, the case since the substitution of the current equity market price, "implied" firm's assets price and volatility into Merton's formula (along with the specifics of the debt contract and the risk free rate), results in estimated spreads that differ greatly from those observed empirically. In fact, in order to obtain spreads of similar magnitude to the observed values, the volatility of the assets needs to be set to unrealistic values (see, for example: Kim, Ramaswamy, and Sunderasan (1993), and Wei and Gou (1997)). More concerning, the values for the firm's assets and volatility implied from equity prices often disagree with those obtained from bond prices.

8. Mathematically, this can be expressed as

$$\text{Distance to Default} = \frac{(\text{MVA} - \text{Default Point})}{\text{MVA} \cdot V}$$

where V is the volatility of the firm's assets, and Default Point = CL + LTD/2. See Vasicek (1984, 1995), Kealhofer, Kwok and Weng (1998) and Caouette, Altman and Narayanan (1998).

9. See Caouette, Altman and Narayanan (1998), pp. 139-153.

10. There are theoretical reasons for this, as well as empirical support. For a discussion of the theoretical rationale, see Sobehart and Keenan (1999).

11. See Keenan, Hamilton and Berthault (2000).

12. For example, standard residual error tests (such as t-statistics in regression) are not meaningful for determining the significance of the candidate variables in these situations.

13. For an overview of bootstrap methods, see, for example, Efron and Tibshirani (1993). In general, variable selection for non-linear models is a computationally intensive problem (see Appendix A).

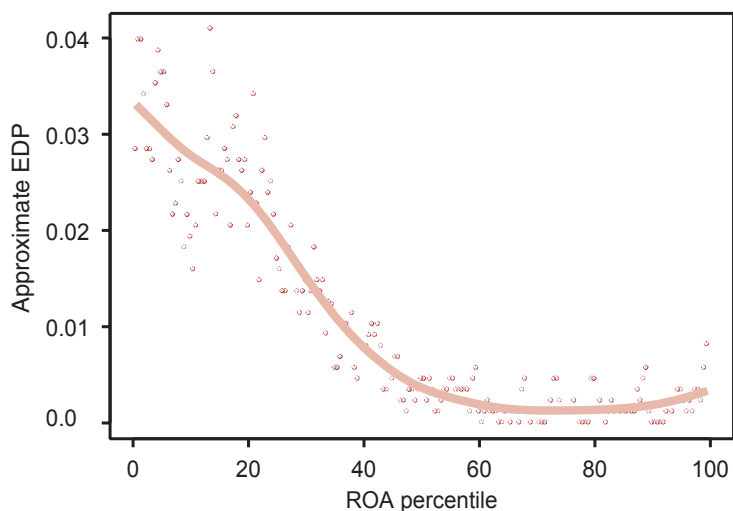
14. The final raw variables are Market Equity, Current Assets, Total Assets, Current Liabilities, Total Liabilities, and Net Income.

Exhibit 3. Key Variables of The Default Prediction Model

Model Variable	Definition ¹⁵	Frequency
Credit Quality	Rating ¹⁶	Credit history when available
Return on Assets	Net Income / Assets	Annual
Firm Size	Log(Assets)	Annual
Operating Liquidity	Working Capital / Assets	Annual
Leverage	Liabilities / Assets	Annual
Market Sensitivity	Stock Price Volatility	Monthly
Equity Growth	Equity Growth Rate	Monthly
Return On Equity	Net Income / Equity	Monthly
Distance to Default	Distance to default	Monthly

It is informative to examine the variables listed in **Exhibit 3** within the context of our model. In order to gain insight into how well these variables differentiate between defaulting and non defaulting firms, and to get a sense of the non-linear nature of the relationships of these variables to default, consider one of the most significant accounting variables in Moody’s model: return on assets (ROA).

Exhibit 4. The Historical Default Rates vs. ROA



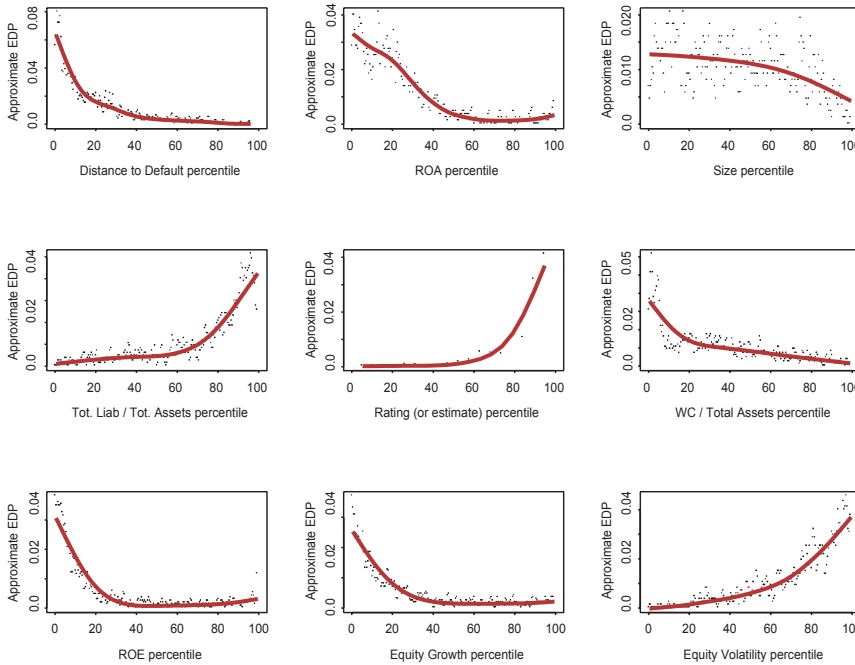
The level of historical defaults versus percentiles of ROA shows both how predictive the variable is (in a univariate sense) and the non-linear nature of the relationship. (Smoothed line superimposed on data.)

Exhibit 4 shows the relationship between the historical frequency of defaults and the value of a firm’s ROA. At each level of percentile of ROA, the percentage of firms with that value that defaulted within one year is shown. The figure shows a clear relationship between level and frequency of default in the historical sample. Note how steadily the observed frequency of default drops as the ROA increases. However, beyond a certain point ROA can be too high, which also may indicate a short-term sales spurt and an inability to sustain performance. Also note the markedly non-linear character of the relationship. Similar relationships can be mapped for each of the input variables used in Moody’s model. These are given, in composite form, in **Exhibit 5**.

¹⁵ Here we define "Assets" as the book value of total assets, and "Liabilities" as the book value of total liabilities.

¹⁶ An estimated rating based on a proprietary rating model is used in place of a Moody’s rating for unrated firms.

Exhibit 5. Relationship between remaining variables and historical default

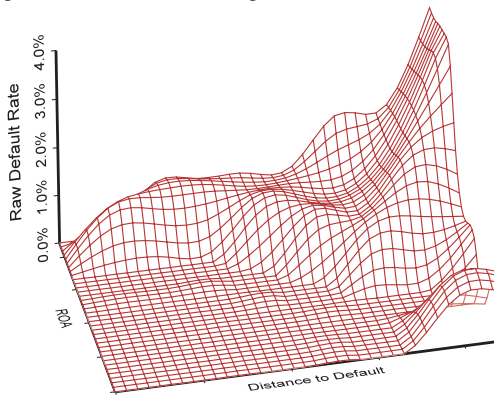


The relative levels of historical defaults versus the observed output of the variable shows both how predictive the variable is (in a univariate sense) and the highly non-linear nature of the relationship. (Loess smoothing used to calculate density.)

Note, however, that the relationships shown in these figures are only univariate in nature. That is, they only show the relationship of a specific variable with respect to historical default frequency *when all other variables are ignored*. The true relationships are, in fact, much more complex, once the additional influences of other variables are included. To give a flavor of this, **Exhibit 6** shows the interaction of two influential variables, in this case the distance to default and a profitability measure, ROA.

Although a clear structure is visible between any two variables, it becomes impossible to visualize in higher dimensions.

Exhibit 6. Frequency Surface Formed By The Interaction Of Two Influential Variables



The relative levels of historical defaults versus the percentiles of the two variables (ROA and Distance to default) shows both how predictive the variables are (in a bivariate sense) and the highly non linear nature of the relationship. The relationships become more “warped” as the dimensionality increases, suggesting potential benefits from using non linear modeling techniques.

Exhibit 6 also gives some evidence of the limitations of a pure market-based approach. The EDP surface behaves differently at different levels of the variable ROA even after accounting for the Merton variable distance to default. This indicates that there is additional structure in the default process not captured by the variable alone. If ROA were not also influential, the surface would be a smooth sheet that followed the curvature of the univariate behavior of distance to default, similar to a sheet shaped like the distance to default plot in **Exhibit 5**.

2.3 A NON-LINEAR MODELING APPROACH

The analysis of variable behavior described in the last section led Moody's to consider using non-linear approaches to modeling credit risk. To accurately capture the high-dimensional and non-linear nature of default events, an approach that was highly flexible was required. One such approach with which Moody's has a long experience is a class of statistical techniques called *artificial neural networks* (ANN). However, traditional applications of ANN technology have involved dauntingly complex network architectures ("black boxes"). Our approach was to develop a type of nested logistic regression that is simple enough to analyze but still captures the non-linear relationships that characterize the default prediction problem.

Because of the flexibility of the approach, and its ability to model high-dimensional and non-linear relationships, Moody's modeling process incorporates extensive sensitivity analysis on models to ensure that they are performing in a rational manner, that their parameters and performance are statistically stable, and that the underlying relationships are well understood.¹⁷ Conceptually the elasticity analysis is similar to the graphical analysis done in **Exhibit 5**. However, for practical purposes, this type of analysis is too informal for model development and we typically use a variety of more rigorous technical methods to conduct the sensitivity analysis.

As a simple example, consider the case of Applied Magnetics Corp., shown at the opening of this *Rating Methodology*. A reasonable question to ask is what are the factors that drive the EDP for that firm? One way to approach this question is to determine the relative influence of different model inputs on the model output.¹⁸ This analysis is straightforward, and is shown in **Exhibit 7**, which uses data taken from the financial statements available prior to January 1998 and January 1999¹⁹. An influence value of zero suggests that the variable is no more or less influential for this firm than it is for the average firm. Values greater than zero indicate variables that adversely affect the evaluation by increasing the EDP, and those less than zero improve it. This influence analysis accounts for components of non-linearity within the model.

From **Exhibit 7**, it becomes clear that the factors driving the model leading up to default were adverse equity price and a general weakening of both market variables and key financial statement variables. Note the dramatic change in variable influence between January 1998 (when the EDP was a little below the population average) and January 1999 when it had a very high EDP. This sensitivity analysis shows that in addition to decreased confidence in the equity markets, the leverage and return on assets also became markedly weaker. In contrast, the 1998 influence analysis shows that the only very strong negative influence was the equity trend, and even that was partially offset by the favorable ROA reported in the previous year.

Exhibit 5, and **Exhibit 6** suggest that the variables that correlate with default tend to be non-linear. While the non-linear nature of the model does make it more difficult to interpret than simple linear ones, this limitation is offset by the increased accuracy in predicting default probabilities that is afforded by a flexible non-linear approach. This is not surprising since traditional credit analysis is a highly non-linear process based on the evaluation of many aspects of credit quality.²⁰

Finally, it is also important to note that, as a result of data limitations, the model parameters must be fit using a data set in which the proportion of defaulting-to-non-defaulting obligors might be different from the actual proportion for the population of public firms. Not all public firms have reliable financial and market information available for parameterizing the model. This is particularly true for defaulting firms whose financial and market information are less likely to be complete or reliable in the time period leading up to default.

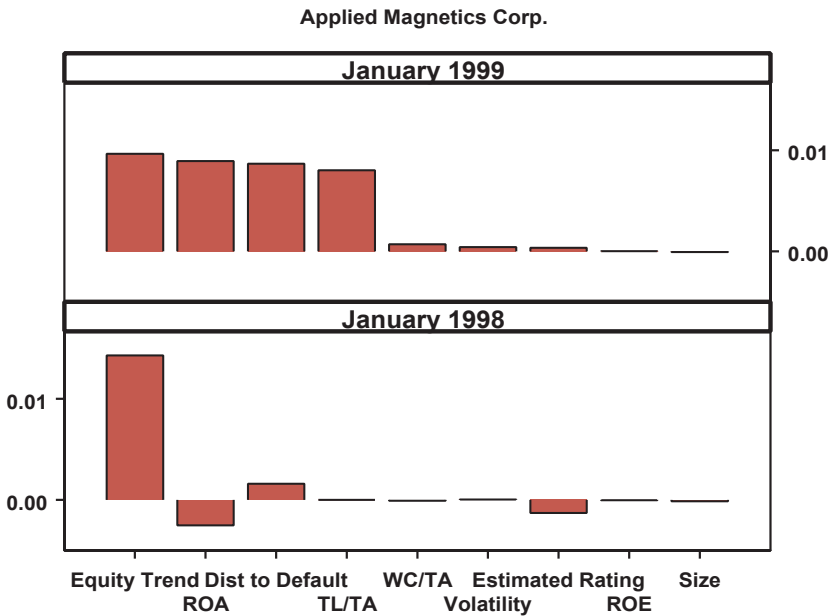
¹⁷ In addition to sensitivity analysis, there are a variety of statistical measures of significance that can be used to determine more conventionally the statistical validity of the model parameters and their dynamics. See Golden (1996) for an overview of these approaches.

¹⁸ For convenience, we define "relative" to be the global influence of the factors beyond the average behavior of the model: $EPD_{[X]} - EPD_{[sector]}$. We also perform extensive local sensitivity analysis.

¹⁹ These financial statements were reported for 1997 and 1998, respectively.

²⁰ See Stimpson (1992) for an overview.

Exhibit 7. An Example Of Influence Analysis of Model Factors: Applied Magnetics Corp.



This figure presents one type of sensitivity analysis that describes the relative influence of each of the variables in the model with respect to the EDP of Applied Magnetics Corp. An influence value of zero suggests that the variable is no more or less influential for this firm than average. Values greater than zero indicate variables that adversely affect the evaluation by increasing the EDP, and those less than zero improve it. This influence analysis accounts for components of non-linearity within the model. Note the dramatic change in variable influence between January 1998, when the EDP was below the population average and January 1999 when it had a very high EDP. On January 7, 2000 Applied Magnetics Corp. filed for protection under Chapter 11.

As a result, the model is parameterized with a reliable sub-sample of the true population. Once we optimize the model parameters, we adjust the probability of default to correct for the fact that this in-sample data set had a slightly different proportion of defaulting-to-non-defaulting obligors (in Bayesian terms, adjust for the priors of the true population). This adjustment is made based on an analysis of a subset of the data in Moody's default database that included over 1,400 non-financial US defaults between 1980 and 1999. (See: *Appendix C: Estimated Default Probability* for a more detailed discussion of the probability adjustments and *Appendix E: Moody's Default Modeling Database* for a description of the our modeling database.)

3. Model Performance And Benchmarking Of Moody's Model Against Other Models

The issue of credit model validation and performance measures can be exceedingly subtle and is often ignored or given cursory treatment in the reported research on credit risk models. To address this, Moody's has recently released a report (Sobehart, Keenan and Stein (2000)) that deals exclusively with the topic of risk model validation and performance analysis²¹. The article describes a comprehensive and rigorous framework that can be used for evaluating certain classes of default models when data are available. The framework specifically addresses validation concerns and provides a methodology for evaluating the performance of wide variety of credit risk models. The article also discusses several alternative measures of model performance, two of which we apply here for illustration purposes.

In the remainder of this section we apply this framework to compare the Public Firm model to other popular credit models. While we briefly highlight some of the features of our validation approach, interested readers are encouraged to review the report and referenced articles as they provide comprehensive discussions of both the general methodological framework and the performance metrics. (See also *Appendix B: An Introduction to Model Performance and Validation*.)

²¹ Also, see Keenan and Sobehart (1999).

We compare several popular univariate and multivariate credit risk models²²:

- (1) Moody's default prediction (non-linear) model;
- (2) variant of the Merton model based on distance to default²³;
- (3) a hazard model based on financial data²⁴
- (4) the original Z-score model (a widely used benchmark);
- (5) reduced Z'-score model (1993)²⁵;
- (6) univariate model based on return on assets (ROA) only.

These models represent a wide range of modeling approaches in order of decreasing complexity. In particular, note that since model (2) is one of the inputs to Moody's Public Firm model (1), any significant discrepancy in the performance of models (1) and (2) shows the degree to which additional information is captured by the use of financial statement data and Moody's ratings.

3.1 MEASURING MODEL ACCURACY

Using an extensive data set for testing the performance of credit risk models represents the opposite extreme from the anecdotal comparisons described in the beginning of this *Rating Methodology*. The performance measures described below provide an accurate accounting of how Moody's model performs over a comprehensive data set of default events obtained from Moody's default database. This database represents the most complete source for public firm default data available.

The most common performance measure used to evaluate credit risk models is the intuitively appealing "number of correct predictions." Default risk models can err in one of two ways. First, the model can indicate low risk when, in fact, the risk is high. Typically referred to as Type I error, this corresponds to highly rated issuers who nevertheless default on financial obligations. Secondly, the model can indicate high risk when, in fact, the risk is low. Typically referred to as Type II error, this corresponds to low-rated firms that should, in fact, be rated higher. It is possible for some credit risk models to commit less of one type of error than another. However, success at minimizing one type of error necessarily comes at the expense of increasing the other type of error. Type I / Type II performance is typically presented in the form of a contingency table or confusion matrix. **Exhibit 8** shows the average confusion matrix produced by Moody's Public Firm model.

The fractions of correct classifications are ordered as the diagonal elements of the array, and the misclassification errors are the off-diagonal elements (Type I and Type II errors). The standard deviations of the misclassification errors are shown in parenthesis. Note that the model produces reasonably low misclassification errors even when the non-defaulting firms in the in-sample data set are about 100 times more numerous than the defaulting firms, which was the case in our sample.

It is often difficult to achieve consistent performance across credit quality when modeling defaults because the infrequent nature of defaults often forces model builders to over represent defaults in the data sets they use to fit their models.²⁶ This results in parameters that do not represent the true default structure. The consistent performance between in and out of sample data sets that Moody's achieves is the result of:

- (a) the extensive use of cross validation that we make during the modeling process (see: *Appendix A: Variable selection and model construction* for a fuller discussion), and
- (b) the probability adjustments to account for differences in the default rate of the data set we use to fit our model and that of the full universe of public firms.

²² We compared several additional models which produced similar results.

²³ A similar adaptation of the Merton model has been popularized by the KMV Corporation. For details see Kealhofer (1999).

²⁴ We implemented the model based on Zmijewski's variables described in Shumway (1998).

²⁵ For a definition of the reduced Z'-scored see Caouette, Altman and Narayanan (1998), p. 121.

²⁶ See Caouette, Altman and Narayanan (1998) for examples.

Exhibit 8. Confusion Matrix For The K-Fold Cross Validation of The Model

	Predicted Non-Defaulters	Predicted Defaulters
In-sample ^{27 28}		
Actual Non-Defaulters	86 % [5 %]	14 % [5 %]
Actual Defaulters	27 % [5 %]	72 % [5 %]
Internal cross-validation sample		
Actual Non-Defaulters	85 % [7 %]	15 % [7 %]
Actual Defaulters	28 % [7 %]	71 % [7 %]
Out of sample / out of time validation		
Actual Non-Defaulters	83% [7 %]	17% [7 %]
Actual Defaulters	26% [8 %]	74% [8 %]

Although Type I and Type II classification errors provide valuable insight, they are limited measures of model performance. In this section we present two simple, yet powerful, comparative performance measures for credit risk models: (1) Cumulative Accuracy Profiles, and (2) Accuracy Ratios.

These metrics are quite general and can be used with many testing schemes to compare different types of models even when the model outputs differ and are difficult to compare directly. Specifically, discrete risk ratings can be compared to continuous numerical outputs. Even the categorical outputs used by rating agencies can be evaluated side by side with “continuous” quantitative model scores. A more detailed description of these measures, as well as additional useful metrics based on the information content of the models, is given in Keenan and Sobehart (1999) and Sobehart, Keenan, and Stein (2000).

3.1.1 Cumulative Accuracy Profiles (CAPs)

Moody’s uses Cumulative Accuracy Profiles (CAP) plots, to make visual, qualitative assessments of model performance. A CAP curve²⁹ is constructed by calculating the percentage $y(x)$ of the defaulters whose risk score is equal to or lower than the one for a fraction x of the population. The “higher” (closer to the northwest corner) the curve the better. **Exhibit 9** shows the Type I CAP curves for several models using the validation sample (out-of-sample and out-of-time). Similar results are obtained for the in-sample tests. Note from the validation results, that Moody’s Public Firm model dominates the other models tested. In particular, it out-performs the modified version of the Merton model.³⁰

²⁷ In-sample refers to observations used for building the model. A fraction of the in-sample set is used to find the parameters of the model, and the rest is used for internal validation checking the model consistency (cross-validation sample).

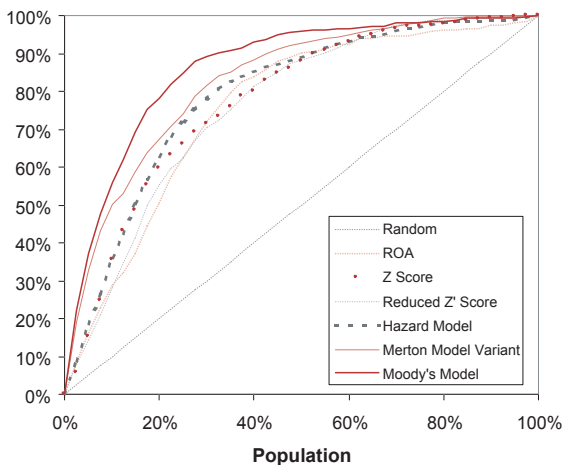
²⁸ We estimated the parameters of our model with a representative sample of the actual population as opposed to the standard procedure in which a 50-50% mix of defaulters and non-defaulters are trained and then the output adjusted with the default rate of the population. The later approach does not take advantage of all the information available introducing selection bias.

²⁹ While similar tools exist under a variety of different names (lift-curves, dubbed-curves, receiver-operator curves, power curves, etc.), Moody’s use of the term CAP refers specifically to the case where the curve represents the cumulative probability over the entire population, as opposed to the non-defaulting population only. In statistical terms, the CAP curve represents the cumulative probability distribution of default events for different percentiles of the risk score scale. See Keenan and Sobehart (1999) and Herrity, Keenan, Sobehart, Carty and Falkenstein (1999).

³⁰ Also note that the performance of this model for the in-sample and validation data sets is in reasonable agreement with the reported performance obtained from similar market-based models. See, for example, McQuown (1993) and Kealhofer (1999).

Exhibit 9. CAP Curves For The Tested Models

Selected CAP Curves



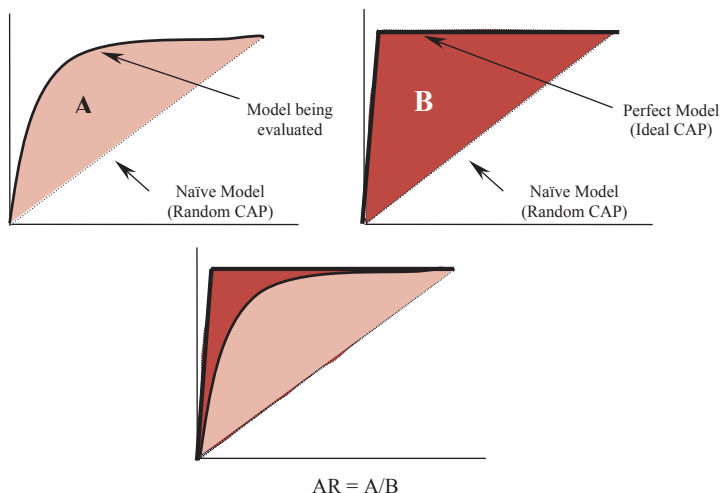
This figure shows the CAP curves for the selected models. All models were tested on the same validation data set. The 45° dashed gray line represents the naïve case (which is equivalent to a random assignment of scores). All models perform considerably better than the random case. Note that the Merton model variant performs almost as well as the nonlinear model in the case of extremely poor quality firms. However, the nonlinear model clearly performs better beyond about the bottom 10% of the populations and is much better at discriminating defaults in the middle ranges of credits.

3.1.2 ACCURACY RATIO (AR)

It is convenient to have a single measure that summarizes the predictive accuracy of each risk measure for both Type I and Type II errors. We obtain such a measure by comparing the CAP of any risk measure with both the ideal (perfect prediction) and non-informative (random) CAPs. The closer the CAP is to its ideal, the more area there is between it and the random CAP. The largest area that can possibly be enclosed is identified by the ideal CAP. The ratio of the area between a risk measure's CAP and the random CAP to the area between the ideal CAP and the random CAP is the Accuracy Ratio (AR). The AR is a fraction between 0 and 1. Risk measures with ARs close to 0 display little advantage over a random assignment of risk scores, while those with ARs near 1 display almost perfect foresight.

The accuracy ratio can be envisioned as the ratio of the shaded region in the graph on the left of Exhibit 10 to the shaded region on the right of the figure. The difference is shown in the bottom of Exhibit 10 below.

Exhibit 10. Heuristic Representation Of The Accuracy Ratio



$$AR = A/B$$

The accuracy ratio is the ratio of (A) the performance improvement over the naïve model of the model being evaluated to (B) the performance improvement over the naïve model of the Perfect Model. It can be envisioned as the ratio of the shaded region in the graph on the left of to the shaded region on the right. The difference of the areas is shown in the bottom of the figure.

Most of the models we tested had ARs in the range of 50% to 75% for out-of-sample and out-of-time validation and in-sample tests for all rated and unrated non-financial public firms.³¹ Empirically, the accuracy of the estimates deteriorates for small samples. In order to reduce the sensitivity of the AR to outliers and the rare-event nature of defaults (small samples) we perform sensitivity tests using bootstrapping.³²

In a loose sense, AR is similar to a Kolmogorov-Smirnov (KS) test designed to verify if the model is better than a random assignment of EDPs. However, AR measures the global discrepancy between the cumulative distribution functions of defaults while a KS-test focuses on the asymptotic value of the maximum discrepancy. **Exhibit 11** shows AR values for the tested models for in-sample³³ and validation tests (out-of-sample and out-of-time). The typical error bound is 0.02.

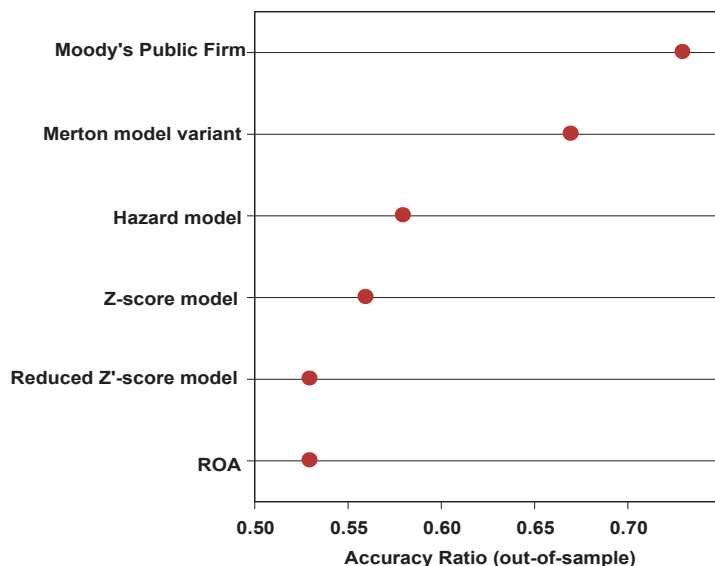
To confirm the validity of the AR figures, we also checked if a particular model differed significantly from the one ranked immediately above it using KS statistics tests over 9,000 independent observations selected from the (out-of-sample/out-of-time) validation set. KS tests support the AR results on the validation sample. More precisely, KS tests showed that only the reduced Z'-score and ROA were not significantly different for this sub-sample. Note, however, that the reduced Z'-score performs significantly better in the in-sample data set.

As in the CAP plots, the Moody's model outperforms the other models. This is natural given the close relationship between CAP plots and AR. There appears to be a significant jump in performance as one moves from pure statistical models to those that include structural information. This can be observed more clearly in **Exhibit 12**, which depicts the same information graphically.

Exhibit 11. Selected Accuracy Ratios

Model	In-sample AR	Validation AR
Moody's Public Firm	0.76	0.73
Merton model variant	0.67	0.67
Hazard model	0.59	0.58
Z-score model	0.56	0.56
Reduced Z'-score model	0.57	0.53
ROA	0.53	0.53

Exhibit 12 Accuracy Ratios For The Models Tested



This dot plot graphically depicts the performance of each of the six default models validated. Note the large gap in performance between the pure statistical models and the two models that incorporate a market based structural model. Also note the large gap between the pure structural model (Merton variant) and Moody's hybrid model. This gap represents the gain in accuracy from incorporating financial statement and rating data into the structural model.

³¹ The accuracy of the models is slightly better for tests performed using samples of rated firms only. However, the validation sample of rated firms is much smaller than the validation sample for all public firms.

³² See Sobehart, Keenan and Stein (2000).

³³ Note that in-sample refers to the data set used to construct Moody's non-linear model.

Exhibit 12 makes more obvious the differences in performance between the pure statistical models (Hazard model, Z-score model, Reduced Z'-score model, ROA) and the two models that incorporate market information (Merton model variant, Moody's Public Firm model). Interestingly, there is also a large gap between the pure structural model (Merton model variant) and Moody's hybrid model. This gap can be seen as representing the gain in accuracy from incorporating financial statement and rating data into the structural model.

It is important to note however, that the Accuracy Ratio is only one measure of model performance. In evaluating our model, we examined a variety of other measures of model performance as well. In some cases, models that seemed similar in performance along one dimension turned out to have quite different performance when measured using other criteria. For example, the simplest non-linear model Moody's examined in building this version of our Public Firm model, a standard logistic regression, performed comparably to the Public Firm Model along some performance dimensions, but was clearly inferior along others such as the information-based measures described in Keenan and Sobehart (1999) and Sobehart, Keenan and Stein (2000). The Public Firm Model described here consistently had the best performance over all of the measures we examined.

4. Conclusions

In constructing a more general class of default prediction models, we considered the value of ratings, equity market data and financial statement information. We constructed a hybrid default risk model based on non-linear regression methods. The model makes use of a small, but carefully selected subset of accounting variables and market data, expressed in terms of key financial ratios and a contingent claims component (expressed as a distance to default). The distance to default is calculated using implied asset values and volatility from market information and our variant of the Merton model.

The use of non-linear modeling techniques allows Moody's to capture complex relationships, the structure of which is estimated from the statistical properties of the default data itself. The difference in performance between Moody's model and a pure contingent claims model indicates the existence of relevant information that cannot be captured using the Merton model variant alone. While market information provides insight, and has proved useful at predicting defaults, it can also reflect changes in investors' preferences not related to the firm's creditworthiness. In contrast, financial statements contain firm specific information which bears directly on the firm's current financial strength and is key for credit risk assessment. However, alone, they appear to be insufficient proxies for credit quality.

Although ratings are not available for most firms because the labor intensive nature of ratings make them economically viable for only a subset of obligors, they provide strong credit signals with respect to long term credit quality.³⁴ As has been shown, the synthesis of ratings with both market and financial information in a hybrid model can lead to models with lower bias and greater overall performance.

We have also attempted to provide a comprehensive comparison between popular credit scoring models using a consistent methodology³⁵ based on their information content. The rigorous benchmarking was performed using data from Moody's extensive proprietary default database of over 1,400 U.S. non-financial corporate defaults. Our results show that Moody's Public Firm model was consistently more effective than the alternative models at predicting defaults within one year, and can be used as an early warning system for credit risk assessment and portfolio risk management.

Moody's research in this area continues actively and we will continue to report on our findings with respect to improved credit risk models, modeling methodologies and validation strategies.

³⁴ See: Keenan, Hamilton and Berthault (2000).

³⁵ See: Keenan and Sobehart (1999) and Sobehart, Keenan and Stein (2000).

5. References

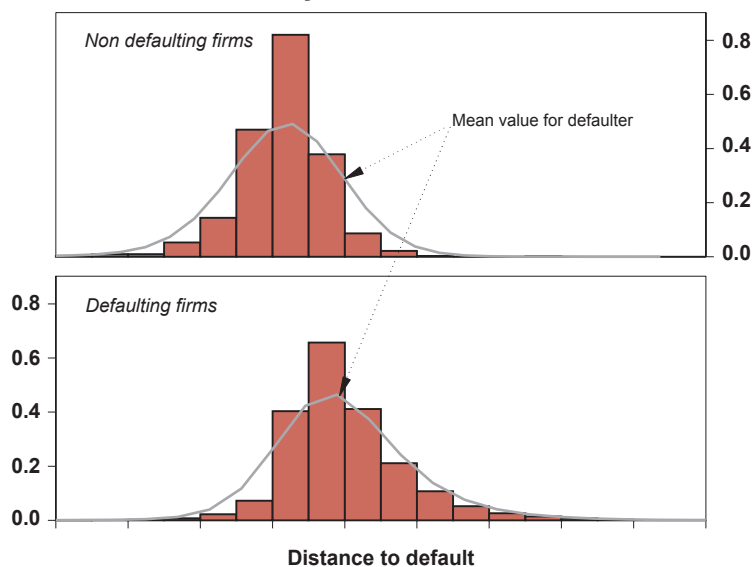
- Altman, E. I.,
“Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy”, *Journal of Finance* (September, 1968), 589-609.
- Altman, E.I.,
Corporate Financial Distress and Bankruptcy (John Wiley, NY, 1993), pp. 184-206.
- Black, F., and Scholes, M.,
“The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81 (1973), 637-659.
- Caouette, J. B., Altman, E. I., Narayanan, P.,
Managing Credit Risk (John Wiley and Sons, NY, 1998).
- Dahr, V. and Stein R.,
Seven Methods For Transforming Corporate Data Into Business Intelligence (Prentice-Hall, NJ, 1997).
- Efron, B., and Tibshirani, R.J.,
An Introduction to the Bootstrap (Chapman, NY, 1993).
- Foster, G.,
Financial Statement Analysis (Prentice-Hall, NJ, 1986), pp. 533-571.
- Golden, R. M.,
Mathematical Methods for Neural Network Analysis and Design, (MIT, 1996).
- Goonatilake, S., and Treleaven, P.,
Intelligent Systems for Finance and Business (John Wiley and Sons, Chichester, UK, 1995).
- Herrity, J., Keenan, S.C., Sobehart, J.R., Carty, L.V., Falkenstein, E.,
Measuring Private Firm Default Risk, Moody’s Investors Service Special Comment (June, 1999).
- Kealhofer, S.,
“Credit Risk and Risk Management” AIMR (1999), 80-94.
- Kealhofer, S., Kwok, S. and Weng, W.,
“Uses and Abuses of Bond Default Rates”, *CreditMetrics Monitor* (First Quarter, 1998), 37-55.
- Keenan, S. C., Hamilton, D. T., and Berthault, A.
“Historical Default Rates of Corporate Bond Issuers, 1920-1999”, Moody’s Investors Service, Special Comment (January 2000)
- Keenan, S. C., and Sobehart, J.R.,
“Performance Measures for Credit Risk Models”, Moody’s Investors Service, Risk Management Services, Research Report 10-10-99 (1999).
- Keenan S.C., Sobehart, J.R., and Hamilton D.T.,
“Forecasting Moody’s Issuer-based Trailing Twelve-Month Default Rate”, Moody’s Investors Service Special Comment (August, 1999).
- Kim, J., Ramaswamy, K., and Sunderasan, S.,
“Does Default Risk in Coupons Affect the Valuation of Corporate Bonds? A Contingent Claims Model,” *Financial Management* (1993), 117-131.
- McQuown, J.A.,
“A Comment on Market vs. Accounting Based Measures of Default Risk”, KMV Corporation (1993).
- Merton, R. C.,
“Theory of Rational Option Pricing”, *Bell Journal of Economics and Management Science*, 4 (1973), 141-183.
- Merton, R. C.,
“On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”, *Journal of Finance* 29 (1974), 449-470.

- Smolensky, P., Mozer, M.C., and Rumelhart, D. E.,
Mathematical Perspectives on Neural Networks (Lawrence Erlbaum Associates, NJ, 1996).
- Shumway, T.,
“Forecasting Bankruptcy More Accurately: A Simple Hazard Model”, University of Michigan Business School working paper (1998).
- Sobehart, J. R., and Keenan, S., C.,
“Equity Market Value and Its Importance for Credit Analysis; Facts and Fiction”, Moody’s Investors Service, Risk Management Services, Research Report 2-7-7-99 (1999).
- Sobehart, J.R., Keenan, S.C., and Stein, R.,
“Benchmarking quantitative default risk models: A Validation Methodology”, Risk Management Services, Rating Methodology, (March, 2000).
- Stein, R.,
“Report on a non-linear rating estimation model”, Internal Moody’s memorandum, (not distributed) (1991). This memo reported on research done to develop a variety of internal credit models firmwide.
- Stein, R., Schocken, S. and Dhar V.,
“A Practical Methodology for Applying Neural Networks to Business Decision Problems”, *Encyclopedia of Computer Science and Technology*, Vol. 38 (1998), Marcel Dekker Publishing.
- Stimpson, D., (ed.),
Global Credit Analysis (Moody’s Investors Service, IFR Publishing Ltd., London, 1992)
- Vasicek, O. A.,
“Credit Valuation”, KMV Corporation (1984).
- Vasicek, O. A.,
“EDF and Corporate Bond Pricing”, KMV Corporation (1995).
- Wei, D. G., and Guo, D.,
“Pricing Risky Debt: An Empirical Comparison of the Longstaff and Schwartz, and Merton Models”, *The Journal of Fixed Income* 7 (1997), 8-28.

6.1 APPENDIX A: VARIABLE SELECTION AND MODEL CONSTRUCTION

Here we discuss the approach used to determine the model structure and the key input variables. In general, variable selection for non-linear (or linear) models is a combinatorial problem. If there are k potential variables, then there are of the order of 2^k possible non-empty subsets of these variables. The variable selection problem is a member of the class of problems known as non-deterministic polynomial complete (*NP*-complete). For any reasonably sized problem the number of possible combinations explodes quickly. For example, with only 20 candidate variables, there are over a million possible subsets. Furthermore, even if exhaustive search were feasible logistically, it greatly increases the chances of obtaining spurious (invalid) statistical results. Thus, a judgmental (non-exhaustive) search must be used to prune the search space initially, based on knowledge of the credit process and constraints on the availability of accounting and market data. Once the minimum set of input variables are selected the model parameters are optimized.

Exhibit 13. Discriminatory Power of the Variable Distance to Default



To do this, we begin with the set of variables found useful by other researchers and by Moody's analysts. We use standard statistical techniques to identify variables that are highly significant in discriminating defaulting firms from non-defaulting firms. We do this in both the univariate and multivariate context. For example, **Exhibit 13** shows graphically the difference in distribution for defaulters (bottom) and non-defaulters (top) for the variable distance to default.

The univariate tests were done in conjunction with an examination of the multivariate structure of the data using linear and non-linear discriminant analysis. The most useful subset of variables, all of which were statistically significant, were then examined in more detail. While all variables were significant, there was still a large number of candidate variables to be pruned. This final step is done by conducting sensitivity analysis on the model performance while selectively including or excluding specific variables.

Because the model is data driven, its parameters can be sensitive to the particular choice of the data sample used for optimization. To obtain robust parameters, the model is re-optimized many times with different subsets of the in-sample data (*cross-validation*) for a given set of variables. (Note that these data do not include data from the hold-out (validation) samples.)

The final set of variables used, the degree of optimization and the model architecture are determined by ensuring that total residual forecast error is within a tolerable range and the overall performance of the model is stable across all subsets of the in-sample data set. This methodology is often referred to in the statistical literature as a k -fold cross validation and we use it extensively in our work. This technique is robust and reduces the dependence effects of the sample.

6.2 APPENDIX B: AN INTRODUCTION TO MODEL PERFORMANCE AND VALIDATION

In this section, we present a brief discussion of the approach that we used to conduct the benchmarking described in Section 3. Interested readers are directed to Keenan and Sobehart (1999) and Sobehart, Keenan, and Stein (2000) for a more detailed discussion.

Designing appropriate tests for benchmarking credit risk models is a difficult task. The lack of data and the averaging effect of long time horizons (spanning over multiple credit cycles) present a challenge to assessing the accuracy and reliability of credit risk models. Transparency in the area of model validation and benchmarking is extremely important since poorly performing models could have serious consequences for risk management practices.

Our validation methodology was developed since many useful measures of system performance are fundamentally statistical in nature. As a result, they are subject to sample variations, which can create the illusion of the superiority of one model over another where no real difference exists. Conversely, model validation methodologies based on predicted distress events, such as the classification error rate, can be problematic since statistical tests for samples with low default rates have extremely low power³⁶ to detect poorly performing models and, consequently, require many default events to produce reliable results. In order to avoid sampling bias when we calculate performance measures, Moody's performs a type of out of sample testing known as "walk forward" analysis, which we describe more fully below.

Because default events are rare and risk scores for consecutive years are highly correlated, it is often impractical to create a model using one data set, and then test it on a separate "hold-out" data set composed of completely independent cross-sectional data. Such out-of-sample and out-of-time tests³⁷ would be the best way to compare models' performance if default data were widely available. However, there is rarely enough default information to support these tests; too many defaulters left out of the in-sample data set will seriously impair the model estimation while too many defaulters left out of the validation sample will reduce the power of the out-of-sample tests.

Our approach is instead to "rationalize" the default experience of our sample by combining out-of-time and out-of-sample tests using walk-forward analysis. The procedure is as follows. We select a year, for example, 1989. We estimate the parameters of the model using all the data available on or before 1989. We then generate the EDP for all the firms available during the following year: 1990. Note that the predicted EDPs for 1990 are out-of-time for firms existing in previous years, and out-of-sample for all the firms whose data become available after 1989. The process is repeated "walking forward in time," using data for every year from 1989 to 1999, and collecting all the out-of-sample and out-of-time observations in a validation set used to analyze the performance of the model. These observations are then used to calculate a variety of robust performance statistics.

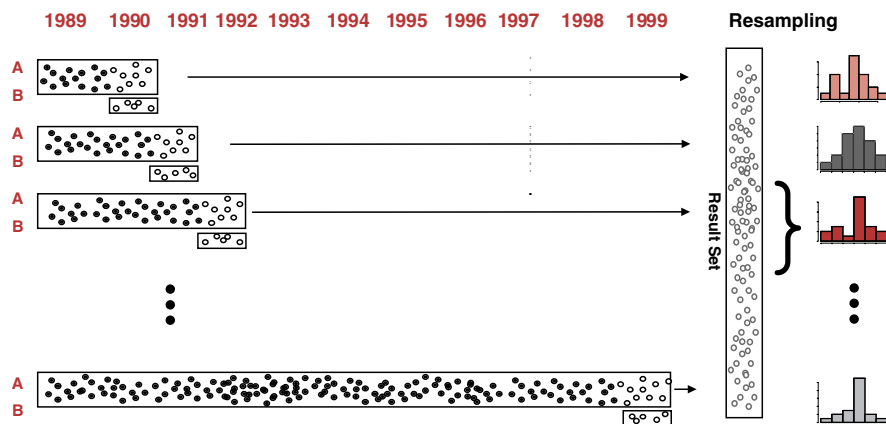
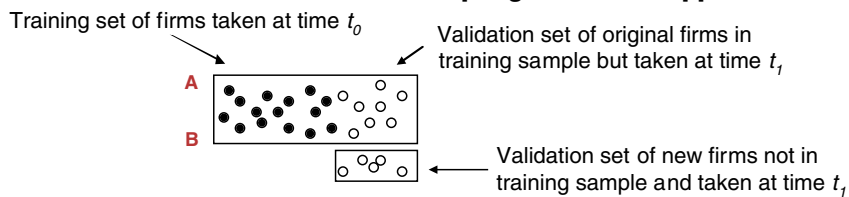
A schematic of the process is shown in **Exhibit 14**. In the exhibit, set A represents the firms that were used to parameterize the model and set B represents firms that were not. The black circles represent the data that were actually used to parameterize the model, while white circles represent data that were held out for validation.

We selected 1989 as the starting year to construct the validation set because our analysis showed that there was not enough data to build a reliable hybrid model with earlier data. The final validation data set contains over 54,000 observations (firm-years), representing about 9,000 different obligors, and including about 530 default events.

³⁶ Recall that statistical power refers to the probability that a statistical test at a particular significance level will unintentionally confirm the null hypothesis when in fact an effect is present. While significance gives information about Type II error, power gives information on Type I error.

³⁷ Out-of-sample refers to observations for firms that are not included in the sample used to build the model. Out-of-time refers to observations that are not contemporary with the sample used to build the model. See: Sobehart, Keenan, and Stein (2000).

Exhibit 14. A Resampling Validation Approach



Moody's fits a model using a sample of historical data on firms and tests the model using both data on those firms one year later, and using data on new firms one year later (upper portion of exhibit). Dark circles represent in-sample data, used to estimate the model, and white circles represent testing data. We do "walk-forward testing" (bottom left) by fitting the parameters of a model using data through a particular year, and testing on data from the following year, and then inching the whole process forward one year. The results of the testing for each validation year are aggregated and then resampled (lower right) to calculate particular statistics of interest. (Reproduced from Sobehart, Keenan, and Stein (2000)).

6.3 APPENDIX C: ESTIMATED DEFAULT PROBABILITY

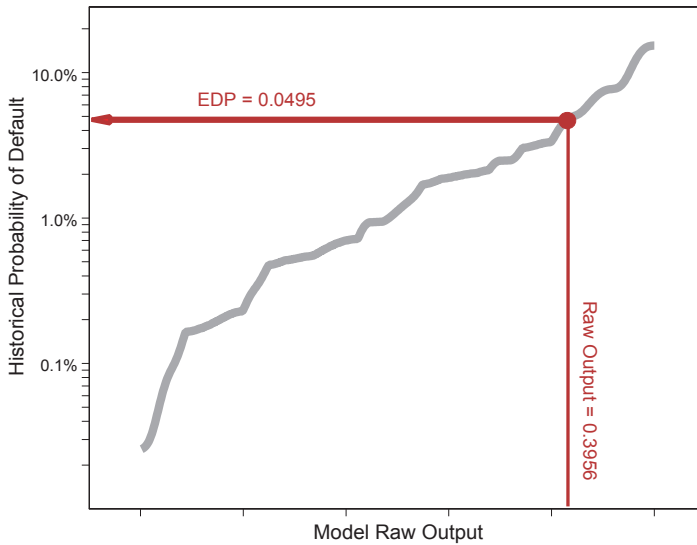
In this section we describe a general, objective and consistent procedure for calculating the estimated default probability (EDP) of an obligor from its financial statements and market information.

It is important to stress that as a result of data limitations, the model must be optimized with a proportion of defaulting-to-non-defaulting obligors different from the actual proportion for the population of public firms. Not all public firms have reliable financial and market information available for use to this end. This is particularly true for defaulting firms whose financial and market information are less likely to be complete or reliable. Thus the model is fitted with a reliable sub-sample of the true population.

Note that even if we could optimize the model to distinguish between defaulting and non-defaulting obligors in exactly the right proportion, the model output would not be the estimated default probability EDP. The output of the model is a score representing different degrees of distress of the firm. Firms tagged as potential defaulters by the model, may or may not default.

In order to convert this score into a consistent default probability estimate, we first need to map the model output to the number of times the model actually predicts the correct distress signal of defaulting firms and the number of times it does not in the in-sample data set. This mapping generates an estimated default probability based on the actual success rate of default predictions after the model is built. This mapping (which is updated every time the model is re-optimized) guarantees that the EDP is historically a correct estimate of the probability of default rather than a simple risk score that happens to be reported as a probability. This approach is similar to the one used in Moody's Default Study. The mapping process is described schematically in **Exhibit 15**.

Exhibit 15 Mapping Model Output to EDPs



Once we optimize the model, we need to correct the probability of default to adjust for the fact that the in-sample has a slightly different proportion of defaulting-to-non-defaulting obligors than the actual population (correct for the prior information). The important point is that as long as the in-sample data set is representative of the true population, the model outputs can be corrected through straightforward probabilistic manipulation. The procedure that Moody's uses to achieve this correction is described below.

At this point it is convenient to introduce the following notation:

1. X_k denotes a set of financial and market information for obligor k . In particular X_k may represent key ratios extracted from the financial statements and capital markets information of obligor k such as the market value of the firm's stock price.
2. $y_D(X_k)$ denotes the value of the default output of the model for obligor k given its financial and market information X_k .
3. n_D and n_N are the number of defaults and non-defaulting observations of the true population of public firms.
4. The default rate of the population is $p_D = n_D / (n_D + n_N)$ and $p_N = 1 - p_D$.
5. $P(D|X_k)$ denotes the probability of default for obligor k given its financial information X_k for the true population. $P(D|X_k)$ is the fundamental quantity we are after. We will use the alternative notation EDP_k to denote $P(D|X_k)$.

Because the final model output is a continuous variable, the historical default rates for each value of $y_D(X_k)$ is calculated using a Gaussian kernel regression method as follows. Let $P_D(y_D(X_k))$ denote the normalized distribution of model outputs for defaulting firms obtained from the kernel regression, and let $P_N(y_D(X_k))$ be the normalized distribution of model outputs for non-defaulting firms in the in-sample data set.

Finally, the estimated default probability for firm k is

$$EDP_k = P(D|X_k) = \frac{p_D P_D(y_D(X_k))}{p_D P_D(y_D(X_k)) + p_N P_N(y_D(X_k))}$$

6.4 APPENDIX D: A NON-LINEAR APPROACH TO MODELING DEFAULT

Moody's uses a wide variety of modeling techniques to develop credit models and continually explores alternative problem formulations and statistical approaches in order to produce the most predictive models possible.

For this version of our Public Firm model, the modeling approach we use can be seen as a higher order logistic regression technique. Instead of creating a single equation, as is done in the traditional logistic model, three equations are created simultaneously. The parameters of these equations are estimated statistically, based on empirical data, using a variant of the familiar optimization technique of gradient descent.³⁸ It is this nested structure, combined with simple non-linear (logistic) transformations that allows the model to map higher order non-linear statistical relationships in the data.

In the current version of Moody's model, the financial inputs to the logistic regression equations are the financial ratios, distance to default, rating, and market variables described in Section 2.2. The output is estimated default probability (or EDP).

Non-linear modeling approaches, such as the one used to parameterize Moody's Public Firm model, have found a variety of applications in the banking and finance domains as well as in a wide range of other commercial areas.³⁹

6.5 APPENDIX E: MOODY'S DEFAULT MODELING DATABASE

In order to construct our models we have created a proprietary database that integrates default information, credit quality opinions, and financial and market data. Our primary sources of default information and ratings information are Moody's proprietary default database and Moody's proprietary ratings database, respectively. We use several third party data sources to obtain relevant financial data. Financial statement data is sourced from the COMPUSTAT U.S. Equities Active and U.S. Equities Research databases. These databases contain both current and historical information on US corporate firms. Market information (weekly prices, and shares outstanding) is collected from IDC.

Because data reliability is key for the development of data-driven models, we frequently verify that the information retrieved from the selected sources is consistent with other sources of fundamental accounting data, such as SEC filings and other sources of market data such as Bloomberg. We also use a series of statistical methods to identify potentially erroneously reported data and fill data gaps. The total number of non-financial U.S. firms used in our database is over 14,000 spanning through the period 1980 to 1999.

For the purpose of this research we defined a default to be the occurrence of any one (or more) of the events listed in **Exhibit 16** for the instruments listed in **Exhibit 17** within one year after the observed financial data. The conditions listed in **Exhibit 16** relate primarily to missed debt payments, distressed exchanges, or formal bankruptcy filings. Note that the exhibit does not exhaustively enumerate all possible cases in which a company may be financially distressed. It does, however, represent a significant fraction of potential distress situations.

The universe of public firms used to estimate and validate our model contains 14,447 public firms with multiple observations for each firm (about 100,000 firm-year observations) and 1,406 default events, of which only 923 have complete financial and market information. The estimated mean default rate of the population of non-financial firms for the period 1980-1999 is about 1.6%.

In **Exhibit 18** we present the industrial composition of the data used in this report. The first column is the industry name. The second and third columns show the total number of companies in each industry and the number of defaulters for which financial statements are available (in some cases a company is included in U.S. Equities-Research because stock prices are available, even though financial data may not be available). Note that this model specifically excludes the financial, banking and insurance sectors. Because of the particular structure of financial institutions, firms in these sectors will be considered in a later extension of the model.

³⁸ For a discussion on the correspondence between maximum likelihood estimation and this gradient descent algorithm, see Smolensky, Mozer and Rumelhart (1996).

³⁹ See Goonatilake, and Treleaven (1995), Dhar and Stein (1997), and Stein, Schoken and Dhar (1998) for an extensive listing of applications in business and finance and for discussions of these applications.

Exhibit 16. Default Types

Default Types
Bankruptcy
Chapter 11
Distressed exchange
Indenture modified
Dividend omission
Missed principal and/or interest payments

Exhibit 17. Debt Type

Debt Types
Bank Loans
Commercial Paper
General Liabilities
Long-term Public Debt
Private Placements
Trade Claims

Exhibit 18. Detail Of The Industries Included In The Construction Of The Model Discussed Here

Industry	Firms	Defaulters
Aerospace and Defense	89	5
Automobile	309	31
Beverage, Food and Tobacco	388	27
Broadcasting and Entertainment	574	62
Building and Real State	759	95
Cargo Transport	290	30
Chemicals, Plastics and Rubber	429	33
Containers, Packing and Glass	74	5
Natural Resources, Precious Metals and Minerals	175	19
Diversified / Conglomerated Manufacturing	85	6
Diversified / Conglomerated Services	567	35
Ecological	57	7
Electronics and High Technology	3,031	234
Farming and Agriculture	144	11
Grocery	206	15
Healthcare, education and Childcare	1,735	124
Furnishings and Durable Consumer Products	327	37
Hotels, Motels, Inns and Gaming	101	9
Leisure and Entertainment	344	41
Machinery (except Agriculture, Construction and Electronics)	435	31
Mining, Steel and Non-precious Metals	425	49
Oil and Gas	838	125
Non-durable Consumer Products	153	11
Personal Transportation	199	34
Food, Personal and Misc. Services	411	58
Printing and Publishing	417	33
Wholesale and Retail Stores	745	150
Telecommunications	338	18
Textile and Leather	387	58
Utilities	312	12
Total	14,447	1,406

**Moody's Public Firm Risk Model:
A Hybrid Approach To Modeling Short Term Default Risk**

Rating Methodology

*To order reprints of this report (100 copies minimum), please call 800.811.6980 toll free in the USA.
Outside the US, please call 1.212.553.1658.*

Report Number: 53853