MOODY'S KMV RISKCALC™ v3.1 MODEL

NEXT-GENERATION TECHNOLOGY FOR PREDICTING PRIVATE FIRM CREDIT RISK

OVERVIEW

AUTHORS

Douglas W. Dwyer
Ahmet E. Kocagil
Roger M. Stein

This white paper outlines the methodology, performance, and key economic benefits of the Moody's KMV Expected Default Frequency™ (EDF™) RiskCalc™ model[1]. A more detailed discussion on the modeling and validation approach can be found in the RiskCalc v3.1 Modeling Methodology document.

The RiskCalc v3.1 model powers the next-generation of default prediction technology for middle market, private firms. With RiskCalc v3.1, Moody's KMV answers an important challenge faced by our customers: "How can we support our decision-making process for extending loans, managing portfolios and pricing debt securities when there is little available market insight into a firm's prospects, as is the case for middle market credits?"

---

[1] For convenience, we use the term "model" in the singular. In fact, RiskCalc is a suite of localized models that share a common framework. Our proprietary database of middle-market financial statement information contains data from each modeling region, which allows us to make modifications with respect to the model inputs and parameters, and to calibrate regional default rates. Each model is adjusted to reflect local economies and reporting standards.

TABLE OF CONTENTS

1 EXECUTIVE SUMMARY

The RiskCalc v3.1 model: The tool that gives you greater insight on risk

The RiskCalc v3.1 model introduces the next-generation default prediction technology for private, middle-market companies. With this model, Moody's KMV continues to innovate, leading the industry with our solution to an important challenge from customers: "How can we improve our decision-making process for extending loans, managing portfolios, and pricing debt securities when there is little available market insight into a firm's prospects, as is the case for middle-market credits?"

With the RiskCalc v3.1 product, Moody's KMV enables you to measure the credit risk of thousands of private companies efficiently and more accurately, expediting underwriting decisions and improving the monitoring of portfolio credit trends. In response to today's increasingly sophisticated credit processes at lenders, and evolving regulations, we have developed a quantitative credit risk model that is intuitive and delivers superior performance.

The increased power of the RiskCalc model to differentiate risk can yield significant increases in profits. Testing by Moody's KMV shows that under simplified but reasonable assumptions:

- A bank using RiskCalc v3.1 might increase the profitability of its loan portfolio by as much as 25 basis points.

- In a competitive environment, for a medium-sized bank, pricing loans with this model would translate into profits of more than $10 million higher on average compared to a competitor that uses a model such as Z-Score. Part of that savings comes from the fact that the bank would also experience a lower default rate.

This white paper describes how we arrived at these figures and explains the underlying research, modeling innovations and key economic benefits of the new RiskCalc v3.1 model framework. Our extensive research combines insights and findings from:

- Best-of-breed modeling approaches built from the insight in Moody's RiskCalc™ v1.0 and the KMV Private Firm Model®.

- Rich data sets available in the world's largest and cleanest private company default database, the Moody's KMV Credit Research Database™ (CRD).

- The introduction of industry sector information, market information, and an adjustment for differences in industry default rates.

The RiskCalc v3.1 model incorporates aspects of both the structural, market-based comparables approach, as used in the original Private Firm Model (in the form of industry-level distance-to-default measures) and the localized financial statement-based approach refined in the original RiskCalc v1.0 technology, now in use by more than 200 institutions worldwide. In this document, we detail how blending market-based (i.e., systematic, sector-based) information with detailed firm-specific financial statement (i.e., idiosyncratic) information yields more accurate models. The enhanced predictive power of the RiskCalc v3.1 model, in turn, yields notably higher and more stable expected profits for lenders and investors.

The result is RiskCalc v3.1: a powerful new model that provides you with a more accurate and comprehensive approach to risk assessment for private firms.

1.1   Strategic Assets of the RiskCalc v3.1 Model

- **Data:** The RiskCalc v3.1 model is based on the richest and cleanest middle-market default data set in the world: the Credit Research Database. This database, gathered from the portfolios of banks and corporate lenders, contains more than 225,000 clean, validated financial statements and almost 4,000 unique confirmed middle-market defaults in North America alone.[2]

- **Extensive validation:** The RiskCalc v3.1 model provides superior predictive results to help you discriminate between subsequently defaulting and non-defaulting firms, and assign accurate probabilities of default to firms. We estimate that the economic impact of our innovations to the RiskCalc v3.1 model can be substantial, possibly running into the millions of dollars for a medium-sized bank.

- **Localization:** RiskCalc v3.1 fits into the consistent framework of RiskCalc models that are estimated individually for each country to reflect the credit and accounting practices of the domicile. This framework allows you to directly compare EDF credit measures worldwide.

- **Support for regulatory terms:** We have designed our new model to meet the requirements of the New Basel Capital Accords, including extensive documentation, validation, and testing.

- **Term structure of default probabilities:** EDF values can now be calculated over horizons ranging from nine months to five years (e.g., 18 months), thus enabling the analysis of nearly any loan-term, investment horizon, or pricing application.

- **Monthly updates:** The higher frequency refresh rate for EDF values allows you to monitor individual credits and portfolios between financial statement reporting periods. Even before risk has shown up in the accounting figures, RiskCalc v3.1 can detect issues that are driving the risk of the industry sector.

- **Credit cycle:** You now have the ability to get either a frequently updated or a relatively stable credit measure by incorporating, or excluding, market trends from the sector.

- **Stress testing:** You have the ability to stress test firms from any point in the economic cycle, including the recent volatile years of 2000-2002.

- **Industry-specific and general credit cycle trends:** Information on credit cycles and industry-specific trends are drawn from the equity market and transformed into credit signals through the pioneering EDF structural methodology of Moody's KMV.

- **Rich set of financial statement factors:** Our new model includes factors such as leverage, profitability, growth, and cash flow to capture the idiosyncratic characteristics of the firm and provide intuitive links for credit analysts.

- **Risk drivers:** You may now conduct detailed analyses of the drivers of a firm's probability of default and analyze factors that represent high potential for increasing risk.

- **Seamless integration into the Moody's KMV credit analytic tools:** including Credit Monitor®, Portfolio Manager™, and Moody's KMV Financial Analyst.

---

[2] These numbers represent twice the volume of data used in RiskCalc v1.0 to predict middle-market defaults in the United States and Canada. This proprietary database, developed and maintained by Moody's KMV, contains more than 6,500,000 financial statements on more than 1,500,000 unique private firms with more than 97,000 default events worldwide. The CRD enjoys similar data richness outside of North America. In Japan, for example, it includes more than 600,00 (cleansed) financial statements and 5,500 (cleansed) defaults.

Moody's | K·M·V

2   THE PRODUCT

2.1   Strategic Innovations from Combining Our Two Powerful Approaches

We designed RiskCalc v3.1 by building upon two of our best products: Moody's RiskCalc v1.0 and the KMV Private Firm Model (PFM).

RiskCalc v3.1 delivers the strategic advantage of blending our strengths into a single product: a new quantitative risk model that responds more quickly to changes in market conditions. After the Spring 2002 merger of Moody's Risk Management Services and KMV, we began to rigorously test our separate, successful approaches and their predictive power for measuring middle-market risk. Our analyses of the model frameworks for Moody's RiskCalc v1.0 and the KMV Private Firm Model revealed the complimentary strengths of each model.[3] Moody's KMV used this innovative research to develop a powerful next-generation model for assessing middle-market credit risk.

The result is the RiskCalc v3.1 model. By incorporating and improving upon the value of both Moody's and KMV's leading products, RiskCalc v3.1 unites risk factors that reflect the important idiosyncrasies of individual firms—quantified as data inputs from financial statements—with systematic market factors. This model also incorporates industry-specific and economy-wide market information, and reflects differences between countries, such as accounting practices, tax codes, and business environments.

Lessons from Moody's RiskCalc v1.0

When we compared the performances of our pre-merger models, we confirmed that the RiskCalc v1.0 model framework is robust and continues to be powerful out-of-sample and across industries. These models were originally introduced by Moody's Risk Management Services and developed as a suite of models specific to individual countries. RiskCalc v1.0 models are currently in use at more than 200 institutions worldwide. Our 2003 research found that firm-specific or "idiosyncratic" factors as measured from financial statement data—in this case, localized models that take advantage of middle-market data—are essential in determining the credit risk of a private firm. This approach represents the primary strength of the RiskCalc v1.0 models, which use non-linear and non-parametric statistical methods to map historical financial statement information to subsequent firm performance. (Please see Section 3, "The Model," for a more detailed description of econometric and statistical estimation techniques.)

Lessons from the KMV Private Firm Model (PFM)

Our 2003 study also confirmed the core insight of KMV's legacy Private Firm Model, which has been embraced by more than 45 leading institutions. We found the PFM comparables model, which uses the distance-to default measure to evaluate credit insight from the equity market, to be predictive as well. Since liquid equity prices do not exist for private companies but are required for a structural approach, PFM used a small subset of financial statement data and a statistical mapping to estimate company value and business risk for its structural model. Moody's KMV testing showed that the PFM approach captures and distills systematic market information about a company's industry that is not fully captured in financial statements alone. (Again, please see Section 3, "The Model," for a more detailed description.)

Today, Moody's KMV is proud to introduce RiskCalc v3.1: a new standard for measuring and predicting credit risk for private, middle-market companies. As we demonstrate in Section 2.2, "Expanded data pool for predictions," the power of this model has been enhanced by further developing the Moody's KMV proprietary Credit Research Database, and by introducing a number of modeling innovations based on Moody's KMV research.

---

[3] See: Stein, Kocagil, Bohn, and Akhavein, 2003.

2.2   Expanded Data Pool for Predictions

In North America, for example, we doubled the RiskCalc private firm dataset for version 3.1, and simultaneously improved data quality by employing advanced statistical techniques.

The performance of the RiskCalc v3.1 model is based in part on an extraordinary and proprietary database developed by Moody's KMV: the Credit Research Database. We have made a significant investment to expand and refine this core data set, increasing its cross-sectional and time series coverage of private firm data. At the same time, Moody's KMV has developed new, cutting-edge processes for cleaning the data and addressing differences in local accounting and reporting practices.

Scope

As of November 2003, the Credit Research Database contained more than 6,500,000 financial statements on more than 1,500,000 unique private firms with more than 97,000 default events worldwide. Our testing has demonstrated that using far richer data and, as a result, estimating more precise model parameters can have a profound effect on the performance of the RiskCalc v3.1 model. Specifically, our ability to control for regional and industry differences improves. (Please see Section 3.4 for a more detailed description of model performance.)

For example, the data set that we used to develop, validate, and calibrate the RiskCalc v3.1 models for the United States and Canada contains more than twice the data used to create the RiskCalc North America v1.0 model. Data in the new model include 112 percent more firms, 95 percent more financial statements and 132 percent more defaults in today's development database, as described in Table 1 below.

TABLE 1   Today's North American RiskCalc v3.1 model predictions are based on more than twice the data and a longer market period than used in RiskCalc v1.0*

U.S. and Canadian Private Firms	RiskCalc v1.0	RiskCalc v3.1	Credit Research Database Growth
Financial statements	115,000+	225,000+	↑ 95 %
Firms	24,000+	51,000+	↑ 112 %
Defaults	1,621	3,764	↑ 132 %
Time period	1989-1999	1989-2002	↑ 3 additional years

* Includes data from both United States and Canada for consistency with the numbers reported for RiskCalc v1.0. Finance, real estate, and insurance companies, as well as not-for-profit and government agencies have been excluded.

Adding data from 2000, 2001 and 2002—a period of intense default activity[4]—is particularly valuable because it extends the database over a complete credit cycle. As a result, RiskCalc v3.1 model users have the ability to use a model calibrated to a wide range of general credit cycle conditions and to stress test the impact of a changing economy on default likelihoods. (Please see Section 2.3, "Support for Regulatory Requirements.")

Data Integrity

In the course of expanding the Credit Research Database, we pioneered numerous processes to improve data quality and integrity. We systematically clean the data to detect obvious data problems or data collection issues, such as whether default information or financial statements are missing for a given institution in a given region or time period, and whether balance sheets balance.

At the same time, we also employ more than 200 specific data-quality metrics and filters, designed in conjunction with participating lenders, to ensure data quality that is essential to model integrity. To support the RiskCalc v3.1 suite

---

[4] The increase in default activity is seen in our private firm database. It can also be seen in bond defaults (Hamilton, 2003) and our proprietary public firm default database.

Moody's | K·M·V

of models specifically, we apply several additional advanced statistical techniques for managing data quality. As a result, the modeling datasets are very clean. For example, in developing the United States model, we removed more than 69 percent of the data submitted by contributors because it contained notable errors. Of particular importance are two techniques we implement to check for misclassified defaults and detect potentially fraudulent statements and/or statements that exhibit data entry errors. (For more information, please see Section 3.4, "Further Modeling Improvements.")

2.3   Support for Regulatory Requirements[5]

We designed RiskCalc v3.1 models to meet New Basel Capital Accord requirements.

The RiskCalc v3.1 model was designed to meet the requirements for default models found in the New Basel Capital Accord (or Basel II) papers. In this section, we describe how our new model supports critical requirements for delivering consistent risk estimates, risk ratings, default probabilities, and model validation. While there is no way to predict the impact that every unexpected event might have on a borrower's loan performance, our model is designed to provide comprehensive guidance to market-based or "systematic" risk factors as well as firm-specific or "idiosyncratic" risk factors that drive the vast majority of credit problems.

Consistent Risk Estimates

The RiskCalc v3.1 model will always produce the same estimate of default risk for a given set of inputs, which meets a critical requirement of the Basel II Accord:

> "The overarching principle behind these requirements is that rating and risk estimation systems . . . provide for a meaningful differentiation of risk, and accurate and consistent quantitative estimates of risk." [Basel II, paragraph 351]

Our model is designed to perform consistently and transparently, which helps support efforts by banks to reliably apply risk assessments across their organizations. As we discuss in the next section, the performance of the RiskCalc v3.1 models is robust and stable. The models provide excellent differentiation between defaulters as well as accurate estimates of default probability. The models are developed using localized subsets of predictive factors. The first generation of our methodology (RiskCalc v1.0) has been established worldwide and is in use at more than 200 banks, corporations, insurance firms, and investment banks.

Forward-looking Risk Ratings

In addition to fundamental financial statement inputs, the RiskCalc v3.1 model incorporates the collective perspective of the market sector in which a firm operates. This is consistent with the Basel II Accord requirement that risk-rating models use all available information in determining a borrower's rating—including the impact of future economic conditions:

> "A borrower rating must represent the bank's assessment of the borrower's ability and willingness to contractually perform despite adverse economic conditions or the occurrence of unexpected events." [Basel II, paragraph 376]

The RiskCalc v3.1 model includes monthly updates with the market's aggregated outlook on the state of the general economy as well as a firm's particular industry. With this design, we leverage indicators that encompass many unexpected events that might affect a borrower's loan performance.

---

[5] Jason Kofman contributed to this section.

Stress Testing Default Probabilities

The RiskCalc v3.1 model is uniquely designed to stress test a firm's sensitivity to the probability of default at different stages of a credit cycle. This feature satisfies a leading imperative of the New Basel Capital Accord:

> *"An IRB (internal ratings-based) bank must have in place sound stress testing processes for use in the assessment of capital adequacy. Stress testing should involve identifying possible events or future changes in economic conditions that could have unfavorable effects on a bank's credit exposures and assessment of the bank's ability to withstand such changes. Examples of scenarios that usefully could be examined are: (i) economic or industry downturns; (ii) market-risk events; and (iii) liquidity conditions." [Basel II, paragraph 396]*

The stress test capabilities of the RiskCalc v3.1 model do more than merely review the historical time series of expected default frequencies for a firm. A historical time series simply restates how the fortunes of a firm changed as the economy and a firm's financial structure changed during one particular historical period.

Our new model allows you to test how a firm, *as it exists today*, would have performed during economic conditions that occurred during, for example, the volatility jump of 1998-1999. In other words, our new model allows you to compare a firm's current probability of default under current market conditions with both worst-case and best-case probabilities of default over the past credit cycle, given the firm's current financial state. This perspective helps you separate the impact of systematic risk from idiosyncratic or firm-specific risk.

For example, Figure 1 shows the impact of general credit cycle conditions over time on a firm's EDF credit measure while holding its financial statement information constant.

FIGURE 1   Stress testing a firm: A firm's Expected Default Frequency™ over time *holding constant* financial statement information
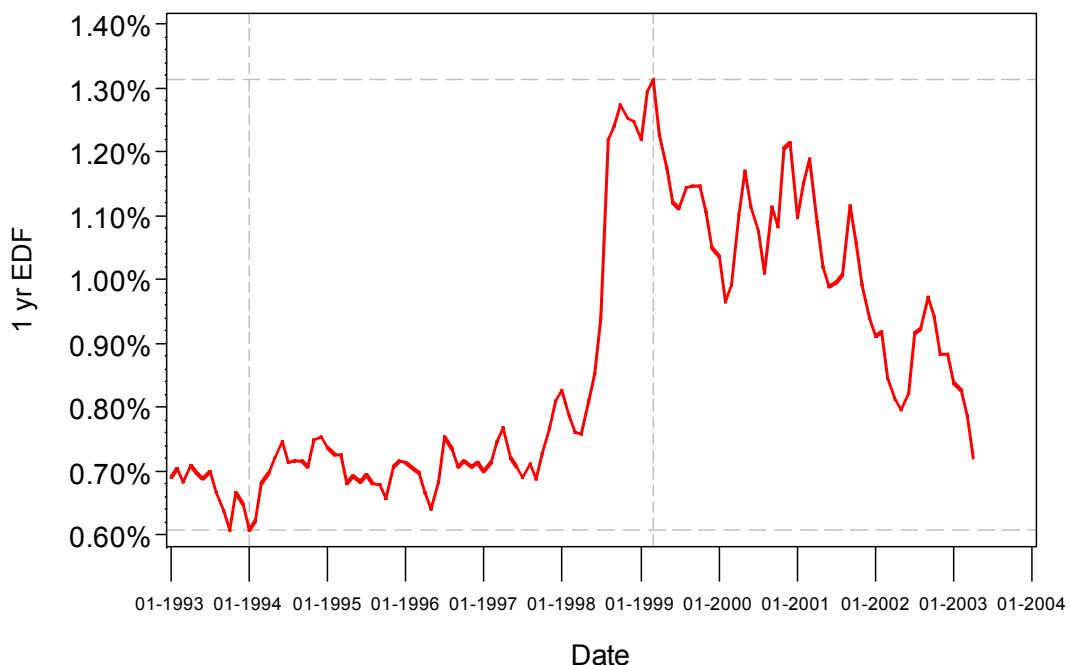
Figure 1 demonstrates how the RiskCalc v3.1 model may be used to compute a firm's best- and worst-case default scenario, given the general credit cycle conditions on a given date. This firm's best-case scenario is an EDF value of 0.60 percent at the beginning of 1994. The worst case-scenario is an EDF value of 1.31 percent in March of 1999, the height of default risk. By May of 2003, the EDF value for the firm is 0.72 percent, near the best-case scenario.

How might a firm have fared over the volatile ten-year period between January 1993 and January 2004? In Figure 1, we stress-tested a firm's performance using its current financial statements and allowing general credit cycle factors to mirror this decade.

As a result, the firm's EDF credit measure is relatively low in January of 1994, when average equity indices and healthy balance sheets were trending upward in anticipation that firms in general—and this sector in particular—had positive future prospects. At the end of 1998 and into 2001, however, the EDF level jumps dramatically to reflect an increase in the market's view of business risk. Importantly, common economic indicators such as equity indices alone did *not* reflect this jump in default risk. Note that the stock market continued to post gains. In other words, even though the stock market was continuing to rise, the Moody's KMV distance-to-default measure already had begun to indicate that the firm's risk of default was increasing because of increased business risk and increased leverage.

The market value declines in 2000 brought asset values down and further drove up credit risk, reducing the distance-to-default. Consequently, given its current financial condition, the firm's EDF credit measure would have remained high for some time and would only recently have begun to return to 1994-97 levels as the market's outlook for the sector improved (or, as the firm's own fundamentals improved if we did not hold them constant in this scenario).

In this example, you can see how a user would calculate the probability of default for this firm at different points in the credit cycle, such as in 1999, when default risk was at its peak. Likewise, the model can demonstrate what the EDF value would be when default rates were bottoming out at 1994 levels.

Validation

RiskCalc v3.1 models are designed to meet the Basel II Accord's stringent requirements for validating ratings:

> *"Banks must have a robust system in place to validate the accuracy and consistency of rating systems, processes, and the estimation of all relevant risk components." [Basel II, paragraph 463]*

Moody's KMV has pioneered the use of empirical validation in commercial credit models. We validate our models using a rigorous testing process that demonstrates their power outside the development sample. These tests include of out-of-sample testing (using defaults and non-defaults which were not used in the model development, such as a "hold-out" sample) and comparisons to other models. For example, the dataset used to validate RiskCalc v3.1 for the United States consisted of 24,768 financial statements from firms that did not appear in the development sample: 23,169 statements from non-defaulted firms and 1,617 statements from 520 defaulted firms.[6]

We find it useful to consider testing two separate but related aspects of model performance: the model's **power**, or ability to rank-order firms from more to less risky, and the model's **calibration**, or EDF level for a group of firms.

The greatest contribution to profitability, efficiency, and reduced losses comes from the model's powerful ability to rank-order firms by riskiness so that the bank can eliminate high risk prospects, allocate analytic resources, and adjust the frequency and resources for monitoring the exposures that have the greatest portfolio impact. RiskCalc 3.1 improves significantly on this predictive power, as detailed in Section 4 on Validation. Moreover, our testing demonstrates that RiskCalc v3.1 does a good job of predicting actual default rates. With the inclusion of sector data and default rates and the inclusion of market data, the default probabilities are more accurate and move more responsively to reflect the change in default rates as conditions change.

---

[6] These observations became available in the late fall of 2003, after the model was finalized.

3   THE MODEL

We redesigned the framework to achieve the superior predictive power of the RiskCalc v3.1 model.

The RiskCalc v3.1 models provide superior predictions of default risk by blending forward-looking systematic information on general and sector-specific credit cycles with a localized approach based on detailed company financial statements. As we describe below, our new model builds on the success of RiskCalc v.1.0 and its firm-specific financial statement model of credit risk by adding equity information, translated into default signals through Moody's KMV structural model framework.

For users who desire a stable estimate of a firm's default risk based only on a firm's financial statements, the RiskCalc v3.1 model can be configured in Financial Statement Only (FSO) mode. In Section 3.1, we describe the FSO mode as well as our process for selecting a limited number of financial ratios to avoid building a model that performs poorly out-of-sample.

If you seek the most accurate determination of the default risk of private company credits and an efficient monthly monitoring process, we recommend the complete version of RiskCalc v3.1. Its cutting-edge, predictive measures deliver accurate EDF levels you can use for origination, pricing, securitization, portfolio analysis, and higher-frequency monitoring. You can also stress test EDF credit measures under different credit cycle scenarios (a Basel II imperative). In Sections 3.2 and 3.3, below, we expand on the challenges of delivering this model, including the value of our distance-to-default calculation, which uniquely equips the RiskCalc v3.1 model to control for industry variation.

3.1   The Financial Statement Only Mode of the RiskCalc v3.1 Model

Obtain stable estimates for default risk using only financial statement inputs.

The Financial Statement Only mode is best suited for users who desire a stable estimate of a firm's default risk for certain applications. The mode includes financial statement variables that capture a firm's long-run performance. Predictions of a middle-market firm's default risk update only as often as the firm updates its financial statements—approximately once a year or, in some cases, once a quarter.[7]

The Financial Statement Only (or FSO) mode is based on financial statement information, similar to Moody's RiskCalc v1.0. In addition, the FSO mode in RiskCalc v3.1 includes industry information.

We selected specific financial ratios to develop a robust and informative model that is transparent, intuitive, and highly predictive for out-of-sample data—even in the presence of some variations in the reported figures due to "creative accounting" procedures.[8]

Ratios

Most standard texts on financial statement analysis discuss ratios that characterize various aspects of a firm's performance. While each of these ratios may provide important alternative perspectives on a firm's condition, our experience is that including a large number of ratios in a quantitative model may yield a model that is "overfitted." In other words, the model will perform very well on the data used to develop the model, but its performance out-of-sample on new borrowers will be poor.

---

[7] Users who desire the most predictive measure of default risk—one that constantly updates and incorporates all relevant data as soon as they are available—will prefer to use the complete version of RiskCalc v3.1 model. For more information, please see Section 3.2, "RiskCalc v3.1, The Complete Version."

[8] We address this issue by implementing non-parametric transformations of the input ratios and combining them in a multivariate context, thus reducing the impact of manipulative noise.

To avoid an "overfitted" model, we developed and refined a process to select a limited number of financial ratios that yield a powerful model. During our selection process, we used statistical tests as well as prior modeling experience to determine which variables to include and exclude from the model.[9]

Our list of financial statement ratios fall under one of the following broad risk factors of financial performance:

- Profitability

- Leverage

- Debt coverage

- Growth variables

- Liquidity

- Activity ratios

- Size

The ratios within each of these groups are viewed as alternative readings of the same underlying construct. In building our financial statement model, we seek to include at least one variable from each group. In the doubling of data and the two years of research that led to this version, we discovered a number of insights that have improved the model. For example, size has typically been found useful as a risk factor. It is true that larger firms default less often, but when we refined our financial statement ratios and included a cash flow ratio, we found that the impact of the size advantage declined in the model. This tells us that a small firm that has very healthy financial statements is not much riskier than a larger firm with comparable financial statements in the Credit Research Database.

- Examples of ratios in the **profitability** group include net income, net income less extraordinary items, EBITDA, EBIT, and operating profit in the numerator; and total assets, tangible assets, fixed assets and sales in the denominator. → *High profitability reduces the probability of default.*

- Examples of ratios in the **leverage (or gearing)** group include liabilities to assets and long-term debt to assets. These ratios measure the size of a firm's debt relative to its assets. → *High leverage increases the probability of default.*

- **Growth variables** are typically sales growth or asset growth. These variables measure the stability of a firm's performance. → *Growth variables behave like a double-edged sword: both rapid growth and rapid decline (negative growth) will tend to increase a firm's default probability.*

- **Liquidity variables** include cash and marketable securities to assets, the current ratio, and the quick ratio. These variables measure the extent to which the firm has liquid assets relative to the size of its liabilities. → *High liquidity reduces the probability of default.*

- **Activity ratios** include inventories to sales and accounts receivable turnover. These ratios measure the extent to which a firm has a substantial portion of assets in accounts that may be of subjective value. For example, a firm with a lot of inventories may not be selling its products and may have to write off these inventories. → *A large stock of inventories relative to sales increases the probability of default; other activity ratios have different relationships to default.*

- **Size variables** include sales and total assets. These variables are converted into a common currency as necessary and then are deflated to a specific base year to ensure comparability (e.g., total assets are measured in 2001 U.S. dollars). The size variable in the U.S. and Canadian v3.1 models is Total Assets. → *Large firms default less often.*

The Appendix provides a listing of the specific ratios used in the RiskCalc v3.1 models for the U.S., Canada, Japan and the United Kingdom. These provide a flavor of how the models are tailored to different lending environments.

---

[9] Each step of this process is described in more detail in the Methodology document. Here we provide only a brief summary.

While each of these ratios relates to varying degrees to credit risk, our research shows a nonlinear relationship between many of these ratios and a firm's probability of default. As demonstrated in Figure 2, below, the probability of default typically decreases as net income to assets (ROA) increases, but the sensitivity of the default likelihood to ROA diminishes as ROA increases.[10] In contrast, we find the impact of growth variables is non-monotonic; for example, both rapid increases and declines in sales are associated with increased default tendencies throughout the world. Both of these observations are quite consistent with the observations of fundamental analysis, and the intuitive nature of the drivers makes the model easier to implement in a credit process.

FIGURE 2   The relationship between a financial statement ratio and default is generally "non-linear"
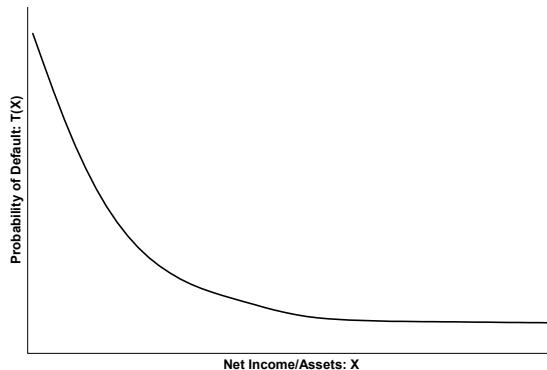
Transformations both illustrate how ratios affect the model and capture nonlinear relationships to produce better predictions.

FSO Functional Form

Our FSO models are based on the following functional form:

$$FSO\ EDF = F\left(\Phi\left(\sum_{i=1}^{N} \beta_i T_i(x_i) + \sum_{j=1}^{K} \gamma_j I_j \right) \right)$$

where $x_1,...,x_N$ are the input ratios; $I_1,...,I_K$ are indicator variables for each of the industry classifications; β and γ are estimated coefficients; Φ is the cumulative normal distribution; F and $T_1,...,T_N$ are non-parametric transforms; and FSO EDF is the financial statement-only EDF credit measure.[11] The Ts capture non-linear impacts of financial ratios on the default likelihood (see Figure 1). We refer to F as the final transform. The final transform captures the non-Gaussian relationship between the default-probability and

$$\left(\sum_{i=1}^{N} \beta_i T_i(x_i) + \sum_{j=1}^{K} \gamma_j I_j \right).$$

This functional form is closely related to a class of models known as generalized additive models. (See Hastie and Tibshirani, 1990; and Pagan and Ullah, 1999.)

This robust model form balances the need to incorporate potential nonlinear behavior with the users' need for transparency. We characterize a model as transparent if it is clear to the user why a change in the input variables resulted in a change in the EDF. For example, if an increase in leverage resulted in a decreased default risk for a corporation, most users would find the implication counter-intuitive. We can easily verify that such nonsensical results will not occur using the FSO mode of RiskCalc v3.1 by examining the transformation for leverage. If leverage is monotonically increasing,

---

[10] It is worth noting that such a relationship cannot be easily captured by simply incorporating quadratic terms (i.e., squares) in a linear regression.

[11] By non-parametric, we mean that the $T(x)$ is a continuous function of x not requiring a specification of a specific closed (or parametric) functional form. We estimate these transforms using a variety of local regression and density estimation techniques.

Moody's | K·M·V

then small increases will always increase a firm's default probability. Verifying that other models share this property—models based on regression trees, neural networks, or linear models with quadratic terms—is feasible but is often not straightforward or intuitive.[12]

3.2   RiskCalc v3.1: The Complete Version

We recommend the complete version of this model if you want the most predictive measures of credit risk for private, middle market companies.

The complete version of the RiskCalc v3.1 model is the most predictive model available for middle market default risk. The model combines forward-looking equity market information that reflects the general credit cycle and the state of the firm's industry with firm-specific data about private companies. This model delivers the following:

- Significantly more accurate EDF levels

- More frequent updates of all relevant information

- The ability to stress test EDF credit measures under different credit cycle scenarios (a Basel II imperative)

After two years of focused middle-market research at Moody's KMV, we determined the need to extend our foundation—the highly effective financial statement model of credit risk established by Risk Calc v1.0—by incorporating systematic risk into the model through market information. As described in Section 1.1 above, our research into the dynamics and drivers of middle-market credit risk revealed the dominant importance of firm-specific or idiosyncratic information in predicting private company defaults.[13] Our testing also suggested the need to support idiosyncratic data with general credit-cycle industry trends that are not included in financial statements, such as systematic risks in the economy such as the indicators picked up in late 1998 by combining equity and balance sheet information (Stein, Kocagil, Bohn and Akhavein, 2003).

To deliver the complimentary insights of these two approaches, our critical challenge was to determine how to blend private firm-specific risk with market insight *despite the lack of market prices for private firms.* The use of equity market information on company prospects and business risk has proven highly successful for assessing the default risk of publicly traded companies. A public firm's stock price can be transformed to indicate the market value of the public firm assets, and thus incorporates the market's perception of both systematic (market) risk and the idiosyncratic (firm-specific) risk that a firm faces.[14]

Distance-To-Default: Using Market Data from Our Public Firm Model to Improve Private Firm Predictions

In contrast to public firms, market prices for claims on the assets of private firms are generally not available.[15] This lack of price-series data makes it difficult to apply directly to private firms the structural approach that has proven so successful for public firms. Our solution—delivered via the RiskCalc v3.1 model—is to begin with an indicator

---

[12] In the course of our research, we also explored more involved modeling approaches but found their impact typically negligible and in some cases detrimental. For example, we examined interactions of the factors (both parametric and non-parametric), alternative modeling frameworks, and alternative "learning algorithms." In addition, such models typically sacrificed transparency and robustness.

[13] Note that this suggests that with appropriate portfolio management, a large portion of the default risk in middle-market portfolios can be diversified away.

[14] For our purposes, a public company is a firm with publicly traded common stock. The stock price of a company is a price for a claim on the firm's assets. For a public firm, Merton (1974) proposed a framework for combining a firm's stock price series with information on the extent of its liabilities to measure its default risk using a structural model framework. This framework has been extended significantly by Moody's KMV (cf., Crosbie, 2003) and it has been shown to be a powerful predictor of default. In fact, this framework has given rise to a new arbitrage strategy that seeks to capitalize on differences in prices between the equity markets and the bond markets. This strategy is often referred to as capital structure arbitrage (cf., Currie and Morris, December 2002).

[15] Some privately traded companies do issue publicly traded bonds. The prices on such bonds are also prices on a claim on the firm's assets. Nevertheless, meaningful price series data on such bonds are rarely available due to the lack of liquidity.

specifically designed to predict the default likelihood of public firms and incorporate this indicator into a model for private firms: the distance-to-default measure.

RiskCalc v3.1 imbeds credit insight from market information at the *industry sector level*, rather than analogizing between a single private firm and implied market values for that company as was done in PFM. That industry sector information comes from the current month information on the sector's average distance-to-default.

The distance-to-default measure used in the Moody's KMV public EDF credit measures represents the number of standard deviations (or distance) between the market value of a firm's assets and its relevant liabilities. This measure combines a firm's liabilities, market value, and volatility of assets into a single measure that determines the probability of default for a public firm (Crosbie, 2003).[16] We found that including the distance-to-default factor, not on the individual private firm but from an aggregation of public companies in the corresponding sector, improves the performance of our private firm models by incorporating forward-looking market price dynamics.

By using the distance-to-default factor, the RiskCalc v3.1 model immediately captures the impact of economic changes that have not yet been reflected in private firm financial statements. Because private firms typically report only one audited annual financial statement per year, information available in these financial statements can significantly lag behind the current state of a company's performance or an industry shift. This lag is exacerbated by the fact that most private firm statements are not available to lenders until three or four months after the statement date for which the firm compiles the data.

We find that changes in the market-based distance-to-default factor for public firms provide a highly predictive leading indicator of the probability of default for similar private firms. In contrast, however, our research results show that the relationship between default behavior and various macroeconomic variables (such as interest rates, GNP, or unemployment rate) that are thought to have an impact in the literature is notably weaker and/or inconsistent over time, making alternative non-market measures of the state of the economy unreliable for default prediction. For example, while default rates in the 1990-1991 and in the 2000-2002 recessions were similarly high, interest rates prior to these recessions were quite different.

When the distance-to-default measure is trending downward, or closer to the default barrier, on average for the public firms in a given sector, we observe that the probability of default for a private firm in that sector should be adjusted upward as this indicator contains information that is not yet in financial statements. As it turns out, empirical evidence supports this view. The distance-to-default variable is critical to the power and precision of the RiskCalc v3.1 model, particularly with regard to industry variation.

Figure 3, below, shows an example of this forward-looking property for an actual private firm that defaulted in 2000. The figure shows the full RiskCalc v3.1 estimate of the probability of default (EDF) as well as the EDF value that would have been obtained using the Financial Statement Only mode that lacks forward-looking factors.

Note how in Figure 3, the EDF generated by the full version of the model (solid line) provides a leading indicator of the increasing risk of the firm in 1998 and 1999. Note that for entire year of 1998, the financial statement mode shows the EDF measure of this firm to be around 6 percent even though the complete model reveals that, in actuality, it has approximately doubled its default probability within the same year. Moreover, the increase in the EDF level that showed up in the financial statements in mid-1999 was predicted by RiskCalc v3.1 well in advance, using forward-looking factors. As the graph reveals, in 1998-2000, the ultimate EDF level was substantially higher using the combined information than it would have been using the financial statements alone.

---

[16] The theory behind this measurement is based on a long tradition of structural models of default that have their origins in the Merton model (cf., Merton, 1974). Recent advances in this line of research are Leland & Toft (1996), Longstaff and Schwartz (1995), and Zhou (1997). This measure has been extensively validated to be a strong predictor of defaults (cf., Kurbat and Korablev, 2002). Moody's KMV currently measures this distance-to-default for every publicly traded firm throughout the world.

Moody's | K·M·V

FIGURE 3   The forward-looking nature of RiskCalc v3.1 versus the Financial Statement Only mode is evident in the case of this private firm that subsequently defaulted in early 2000

LaRoche Industries Inc.
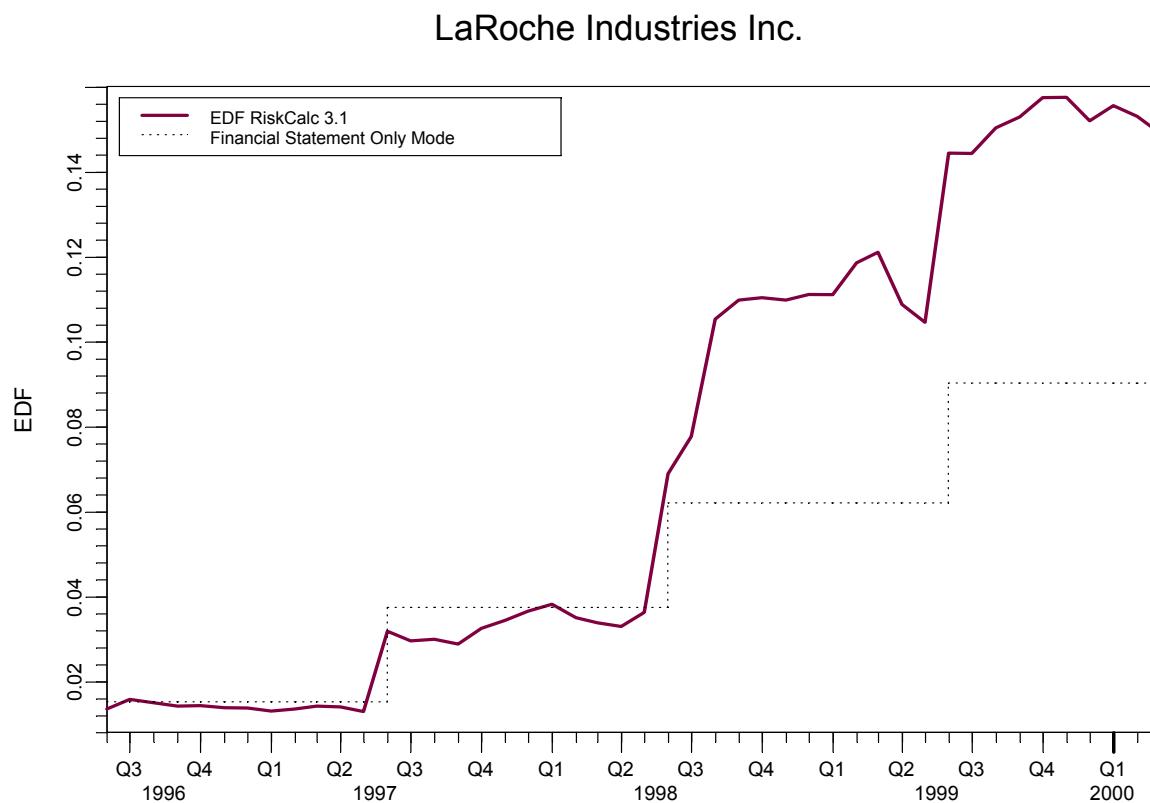
3.3   Introducing Industry Variation to the Model

The RiskCalc v3.1 model introduces the ability to control for industry variation, an important factor in tracking default risk.

When we introduce the distance-to-default factor, industry-wide trends in the public markets are quickly reflected in estimates of private firm default risk. This factor is important when characterizing private firm default risk within industries. By controlling for industry variation, the RiskCalc v3.1 model:

• Corrects for intrinsic differences in default probability across industries

• Adjusts for differences in interpretation of financial ratios across industries, and corrects for spurious effects

• Improves EDF performance and accuracy

Controlling for industry effects yields a modest increase in model predictive power by more accurately ordering firms from more risky to less risky, as demonstrated by a higher Accuracy Ratio. Controlling for industry effects also delivers better accuracy in the probability of default level, as demonstrated by a substantial increase in log likelihood measure. The higher the difference in Log Likelihood, the better the predicted default rates line up with realized default rates. Before we describe in more detail how we tested the final model in Section 4, it is useful to understand some of the intermediate results of our research. Table 2 provides evidence of the importance of capturing industry effects in the FSO mode of RiskCalc v3.1.

TABLE 2   Increase in model power and accuracy from introducing industry controls

	One-year Horizon		Five-year Horizon	
	Accuracy Ratio	**Relative Increase in Log Likelihood**	**Accuracy Ratio**	**Relative Increase in Log Likelihood**
FSO mode without industry controls	54.4%		38.1	
FSO mode with industry controls	55.1%	58.16***	38.8	99.8***

*** Indicates a P-value of less than 0.01 percent.

In this table, and hereafter, Accuracy Ratio (or AR) is our measure of the model's ability to rank order of credits. Increases in log likelihood, on the other hand, measure the extent to which the model's EDF values match observed default rates. For further details, see Dwyer and Stein (2004), *Technical Document on RiskCalc v3.1 Methodology* (Technical Document).

We find that both the power and calibration of default risk prediction improve with the ability to differentiate by industry. This enables the model to incorporate differences in average default rates across industries, and to control for spurious effects between industry and model variables. Lenders have long recognized the importance of industry in analyzing a firm's fundamentals. Model builders have not tackled this issue to date because incorporating industry requires significantly more data than a model without it. A typically limited set of defaults, when divided into industries, can become too small to support building a useful model.

To understand how industry sector impacts results from the Financial Statement Only mode, consider the following example. Both the RiskCalc v1.0 and RiskCalc v3.1 models include inventory-to-sales as a financial ratio. High levels of inventories are consistently associated with high default rates. This ratio is typically valuable because a relatively large stock of inventories may be a signal that a firm is not generating revenue and, as a result, a firm may have to write-off a substantial portion of these inventories. Important industry exceptions do exist, however: some sectors may not accumulate any inventories in the normal course of business. In the services, construction, mining, transportation, utilities, and natural resources sectors, more than 40 percent of these firms do not maintain inventories (see Table 3). Despite an absence of inventory buildup, these firms can still be fairly risky and might not warrant positive "credit" in the risk calculation for low inventories, or, in Version 3.1, for their change in inventory.

By estimating the model with industry-specific adjustments, we control for this issue empirically: we adjust for differences in average default rates across sectors, and at the same time we correct for spurious effects that may be caused by some model variables.

TABLE 3   Percent of observations with zero inventories by sector

Sector	Percent
Agriculture	28.10%
Business Products	9.84%
Communications and Hi-Tech	19.20%
Construction	42.60%
Consumer Products	9.52%
Mining, Transportation, Utilities and Natural Resources	45.00%
Services	49.90%
Trade (Retail & Wholesale)	15.10%
Unassigned	35.90%

3.4 Further Modeling Improvements

We invested in additional research to improve data quality and default risk estimates by the RiskCalc v3.1 model—with valuable results.

Moody's KMV researched a number of additional techniques to address specific challenges we faced when modeling and predicting default risk. While we did not choose to implement all of these techniques in the final RiskCalc v3.1 model, each test confirmed the robustness of the model. Below we provide a detailed analysis of three types of these techniques and our results. For additional technical details, please see the Technical Document. Our research falls into three categories:

1. Managing data quality

2. Alternative estimation techniques

3. Extending the term structure

3.4.1. Managing Data Quality

As described earlier in this paper, the RiskCalc models are estimated based on the Credit Research Database (CRD).[17] Using a larger set of cleaner data by and of itself improves model performance, even without any modeling improvements. To demonstrate the impact, we conducted an experiment in which we re-estimated the U.S. version of RiskCalc v1.0 using exactly the same variable construction on the new data and compared the performance of the re-estimated model to the original model. The re-estimated model outperformed the original model by 3.5 and 5.3 points at the 1-year and 5-year horizons, respectively, out-of-sample. The increase in the likelihood was also dramatic.

In addition to developing and implementing a battery of tests and diagnostic tools to manage data quality, it is instructive to highlight two pioneering techniques we found valuable for managing the effects of misclassification errors and questionable accounting. Both techniques proved useful in the data-cleansing process because they identified issues of integrity that standard methods missed. Both techniques discussed below also helped us better interpret the model.

---

[17] As described above, the CRD contains 6.5 million financial statements on more than 1.5 million unique private firms with more than 97,000 default events worldwide.

Improved data quality leads to improved predictive power

As a result of our extended data acquisition efforts, our data coverage has vastly expanded. From a modeling point of view, an appropriate question to address is whether this expansion in quantity was also coupled by an enhancement in quality of data. Thus, in order to assess whether the new and broader dataset is of better quality, we designed the following experiment:

- We used the variable set of RiskCalc version 1.0 as given (no variable search was performed).

- We re-estimated corresponding transforms and coefficients, which yielded us a "pseudo new" model using pre-2000 observations in our current expanded development data sample.

- We calculated the corresponding accuracy ratio (AR)[18] on the hold-out sample (2000-2002) and compared it with the corresponding AR of RiskCalc version 1.0,[19] which was estimated using data from the same period but prior to introducing advanced data cleansing and adding significant new data to the CRD.

- We performed a log-likelihood test in order to assess whether the precision of the EDF values generated by the pseudo new model are better or worse when compared with RiskCalc v1.0.

Table 4 displays the ARs for the pseudo new model and RiskCalc v1.0 on the same sample for the one-year horizon, and the conclusion is clear: the new data are of higher quality than our original example and yield a better model. As Table 4 suggests, the pseudo model dominates RiskCalc version 1.0 for both time horizons. In addition, we observed that the log likelihood difference, which measures the precision of the levels of EDF measures, is substantial between the two models.

Thus, we conclude that the new data yield a model that dominates the original model both in ranking ability and the precision of calculated EDF measures. This can be interpreted as evidence that the quality of data is superior to the former dataset on which version 1.0 was developed, and that this difference translates into an improvement in the model's predictive power.

TABLE 4   Impact of better data on model performance

	One-year Horizon		Five-year Horizon	
	AR	Relative Increase in Log Likelihood	AR	Relative Increase in Log Likelihood
Original RiskCalc v1.0	48.2%		40.1%	
Re-estimated "RiskCalc v1.0"	51.7%	53.4	45.5%	85.18

The Out-of-Sample Performance of RiskCalc v1.0 and the same model re-estimated using new data — North America

---

[18] For definition of accuracy ratio, please refer to the Appendix in Kocagil and Reyngold (2003).

[19] Since RiskCalc version 1.0 was released in early 2000, the 2000-2002 hold-out sample is out of sample for both RiskCalc v1.0 and the "pseudo new" model.

Moody's | K·M·V

Misclassification Errors

Default prediction challenge: Identifying defaults. The process of matching default events in the CRD with financial statements requires managing the potential for defaults to be misclassified as non-defaults and vice versa. These misclassification errors—which have the potential to compromise model performance—are endemic to middle-market lending institutions and typically occur in the following common scenarios:

- Nearly all CRD participants have transitioned from paper filing systems to electronic databases. Consequently, a complete history of all default and financial statement information is almost never available.

- Many CRD participants have acquired other institutions, and the process of integrating data systems is often incomplete. As a result, we often do not know the exact date of a default—only the date by which the default was reported. Further, we do not always receive all the information regarding accounts that are non-accruals or have been charged off.

- Mapping a subsidiary's default to the financial statements associated with the parent company often creates data integrity issues.

Researching alternatives. We applied a class of techniques developed by Hausman et al. (1998) to detect misclassified information. Hausman's technique is based on examining a firm's classification rates for default probability and investigating deviations from anticipated default behavior. For example, if firms classified as least likely to default demonstrate substantial default rates, this result could be taken as evidence of non-defaults being classified as defaults. Alternatively, if many firms classified as most likely to default do not default, then this result could be taken as evidence of defaults being classified as non-defaults. Hausman et al. present a variety of parametric and non-parametric techniques for sizing the extent of misclassification.

Our results: We performed this analysis across our development sample as a whole. We also examined the issue for each individual institution that contributed data to the CRD. Our analysis confirmed data misclassification errors that were affecting several of the smaller institutions that contribute to the database. One such contributor, for example, stopped providing new data after being acquired by another institution. For this contributor, defaults after the acquisition were not observed and hence were misclassified as non-defaults. In another case, we asked one contributor to investigate a set of defaults among firms that our model considered as very unlikely to default. After reviewing the firms, the contributor determined that the firms did not in fact default and we corrected the data.

Benford's Law

Default prediction challenge: Questionable accounting. Each batch of raw accounting information in our database has the potential for a variety of data integrity issues, including:

- A loss of information can occur when raw data files are read or written if an inappropriate format is used. (Sometimes this loss of information is obvious and sometimes it is not.)

- Information can be lost when variables are standardized and normalized.

- Human error can be introduced when credit analysts enter financial statements into a firm's system.

- Financial statements can include data items that are rough approximations (even guesses) as opposed to true statements of value. Such approximations are thought to be common in unaudited financial statements.

- Financial statements could include outright fraud to avoid taxes or to ensure credit receipt.

Researching alternatives. A new statistical technique has been developed by Nigrini et al. (1996, 2000) to address these types of problems. When financial statement data values are appropriately rescaled, Benford's Law suggests that certain types of figures should follow a specific statistical distribution. Nigrini (2000) advocates using departures from this distribution as a method for signaling potential anomalies in accounting data.

Our results. We applied the Nigrini technique across the different banks in our sample for each of the input variables to our model. We found evidence of excessive rounding in some variables and at certain banks.[20] We experimented with different methods for managing this loss of information. For example, we flagged observations in which there appeared to be a high probability of excessive rounding errors. We then estimated the model with and without these observations and compared model performance both in- and out-of-sample. On balance, we found that model performance was stronger when we included these observations, perhaps because the practice of rounding is a fairly common one in the middle market. Therefore, the final model was estimated with these observations, but we retain the ability to flag potentially suspicious accounting practices.

3.4.2. Alternative Estimation Techniques

The statistical framework we developed and refined for RiskCalc v3.1 modeling balances the need for in-sample predictive power with the need for transparency and out-of-sample performance. In this section, we describe how we tested numerous alternative modeling and econometric techniques to see whether they improved model performance.

Random Effects

Default prediction challenge: Period-to-period dependence. One step in developing the RiskCalc models involves estimating a probit model on a longitudinal (or panel) database. One standard assumption for such a model is an independently and identically distributed error term. Econometric techniques have been developed to allow modelers to relax this assumption in a longitudinal framework.

However, if there is period-to-period dependence, and the assumption of an independently and identically distributed error term is not relaxed, the modeler could inadvertently: (1) include the wrong variables in the model, and (2) produce a model containing inefficient parameter estimates.

Researching alternatives. One of the several factors that govern our decision to include a variable in the RiskCalc v3.1 models is whether or not the variable is statistically significant. The conventional tests for statistical significance in a probit model are not valid *if the error term is time dependent.* Consequently, using these tests by themselves could allow a user to conclude that a variable is statistically significant when it is not. Fortunately, the so-called generalized estimating equations (GEE) yield asymptotically valid hypotheses testing even if time-dependence is incorrectly specified (cf., Zeger and Liang, 1986). To investigate the severity of the assumption violations in our data, we conducted experiments in which we applied this technique. When we did, we found some evidence of dependence. However, while standard errors tended to increase, our conclusions overall remained the same—highly significant variables continued to be highly significant.

Our results. The quality of parameter estimates may be improved by estimating the model with different forms of time dependence in the error term. If no time dependence of the error term is allowed when estimating the model, then the parameter estimates are inefficient and the variance of the estimation error is not minimized. Though this tends to be more of an issue with sparser data sets, we re-estimated both the one-year and five-year models for the United States and Canada, allowing for several different forms of time-dependence in the error term within a firm. We found that the coefficients were stable to different specifications of time dependence in the error term. Further, the rank order of the predicted default probabilities generated by our model was not sensitive to the choice of time-dependence that we used to estimate the model.

Duration Modeling

Default prediction challenge: Default events or intensities and the problem of censorship. To date, RiskCalc private models have been based in part on probit or logit specifications. The dependent variable is taken as a discrete event: whether or not a firm defaults within a certain window of time. We typically estimate separate one-year and five-year models, which allows the relative importance of different variables in predicting default rates to vary at different

---

[20] By excessive rounding, we mean more numbers in which the second and third digit were both zero (e.g., 1,000 or 200) then would be expected according to Benford's Law.

horizons, which we see as an advantage of the approach. A possible shortcoming of this approach, however, is that it does not provide a natural means of dealing with censored observations.[21]

Researching alternatives: Duration approaches directly model the survival time of a firm. Under such models, the dependent variable is the time until default. If a firm has not yet defaulted, then the time to default is taken as the time until it last could have been observed as having defaulted, but it is treated as a censored observation in the likelihood function. A common duration model is the Cox proportional hazards model (for details, see Klein and Moeschberger, 1997). This model allows for the hazard function of a firm to be non-parametric given a set of explanatory variables. A change in the explanatory variables makes a proportionate adjustment to the hazard function. Therefore, the Cox proportional hazards model imposes the restriction that relative importance of different variables for predicting default is constant across all horizons.

Our results: We concluded that model performance was not highly sensitive to specifying the model as either a discrete choice or as a duration model. We also found that the added flexibility associated with estimating two probit models at two different horizons does appear to improve model power.

We performed experiments in which we estimated a duration model using a Cox proportional hazards model as well as discrete choice approaches. We found that the ordering of firms from high-risk to low-risk by the Cox proportional hazards model and our one- and five-year models were very similar (i.e., the rank correlation was high). Interestingly, we found that our one-year model outperformed the duration model at the one-year horizon and that our five-year model outperformed the duration model at the five-year horizon (as measured by AR). As it turned out, the differences in model power were modest. As a result, we consider it reasonable to view default prediction as either event prediction or hazard rate estimation, with the caveat that hazard models are slightly less powerful. Because our formulation permits the inclusion and weighting of different factors at different horizons in a more natural sense, we prefer the discrete choice formulation.

3.4.3. Extending and Filling In the Default Term Structure

Default prediction challenge: Developing a continuous term structure for analyzing obligations of any maturity. Many users are eager to estimate EDF credit measures for the actual term structure, rather than using the one-year and five-year cumulative default probabilities produced by the original RiskCalc model. For example, a bank making a three-year loan is likely to be more interested in the three-year EDF credit measure rather than the one- or five-year EDF credit measure.

Researching alternatives. We found, in our Credit Research Database, confirming evidence that obligors appear to exhibit mean reversion in their credit quality. In other words, good credits today tend to become somewhat worse credits over time and bad credits (conditional upon survival) tend to become better credits over time.[22] Default studies on rated bonds by Moody's Investors Services support this assertion, as do other studies. We found further evidence for mean reversion in both our proprietary public firm default databases and in the Credit Research Database data on private firms.

To test for the presence of mean reversion, we examined a number of approaches to modeling the term structure of default probabilities. Our results showed that we could capture the observed phenomenon of mean reversion in credit risk through a parametric distribution. Under this distribution the hazard rate can either monotonically increase or decrease depending on the selection of the distribution's parameter values. This property of monotonically increasing or decreasing term structures makes the parametric distribution well suited for use in default models.[23]

---

[21] Typically, data sets used for modeling default will have experienced some degree of censorship. If default is defined as a credit event within x years of the date of the observation, then observations that are within x years of the end of the data collection period will not have had a full x years in which to default. A discrete choice model does not provide a direct means of handling such observations.

[22] It has been argued that high quality credits yield positively sloped term structures and low quality credits yield negatively sloped term structures (Bohn, 2000). To the extent that spreads are reflective of credit risk, these findings are broadly consistent with the mean reversion hypothesis.

[23] Hazard rates could potentially be non-monotonic. For example, newly issued debt instruments tend to have low hazard rates initially, which start increasing in subsequent years (Keenan, 1999). Nevertheless, we find that the potential error introduced by the assumption of monotonic hazards is relatively small.

FIGURE 4   Hazard rates implied by Moody's Investors Service Default Studies

Hazard rates as implied by estimated parameters
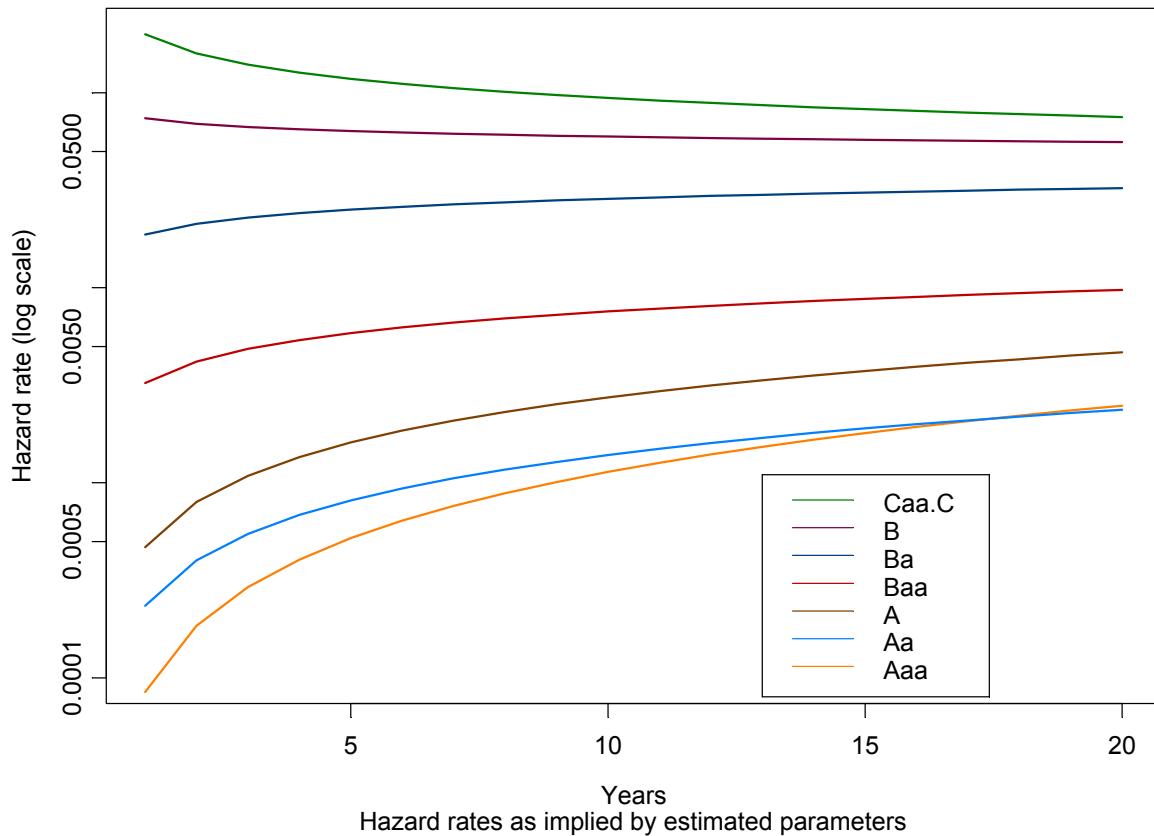
Figure 4 presents the hazard rates (the instantaneous default probability) implied by fitting the parametric distribution to actual survival rates for different rating classifications. The actual survival rates are taken from Exhibit 44 of Hamilton (2003). Note that the high quality credits have increasing hazard rates and that the low quality credits have decreasing hazard rates.

Our results. The new version of RiskCalc term structure framework incorporates mean reversion through the use of a parametric distribution. It gives users more flexibility when modeling probability of default by allowing users to obtain a cumulative default probability for any duration between nine months and five years. The annualized, cumulative, and forward EDF credit measures may be derived from the survival function.

It is important to note that this term structure is different than it would be if estimated using a Cox proportional hazards model. The Cox model imposes a uniform shape of the hazard function across all exposures. As a result, the Cox proportional hazards model, as conventionally estimated, could not capture the mean reversion we observe in actual company data, both public and private.

Accounting for mean reversion is important when pricing loans. For example, if a user assumed that the one-year default probability of a given loan would remain constant over time, the user would under-price higher quality credits since their default probabilities tend to deteriorate over time. Likewise, the user would risk over-pricing low quality credits, as default probabilities for these firms tend to improve over time.

4   MODEL VALIDATION

The superior performance of the new RiskCalc v3.1 is unambiguous. This section presents a number of our analyses.

With the introduction of the RiskCalc v3.1 models, Moody's KMV continues to lead the industry in research, development and application of rigorous, precise default model validation techniques. Our tests indicate that the new RiskCalc v3.1 models outperform all others by measurably large, statistically significant, and economically meaningful margins. As we describe in the sections below on model power and calibration over time, our exhaustive validation assessments demonstrate that:

- RiskCalc v1.0, our first-generation model, continues to perform better in discriminating between more and less risky firms than other available alternatives in a pure out-of-sample context. This model maintained its predictive power during one of the most active periods of default activity since 1920 (Hamilton, 2003).

- The Financial Statement Only version of RiskCalc v3.1 is significantly more powerful than RiskCalc v1.0.

- Our next-generation model, RiskCalc v3.1, is considerably more powerful than RiskCalc v1.0, both in- and out-of-sample, and is substantially more powerful than the alternatives.

- Including our proprietary, forward-looking distance-to-default factor notably improves the performance of the model over shorter horizons when compared with the Financial Statement Only model. As expected, the effect is less pronounced over longer time horizons since the information in the market has time to impact and be reported in the firm's own fundamental accounting data.

- These results were robust out-of-sample and out-of-time in walk-forward and cross-validation studies and on data that arrived after the completion of the model.[24]

Regulators increasingly emphasize the importance of validation. The New Basel Capital Accord calls for additional procedures for testing models, suggesting that banks:

> *"Establish a rigorous statistical process (including out-of-time and out-of-sample performance tests) for validating the selection of explanatory variables; and indicate circumstances under which the model does not work effectively…"*
> *[Basel 2001, Section 252, and Paragraph 599b in the 3rd BIS Consultative Paper]*

It is challenging to test credit processes rigorously due to the relative rarity of default events and the large impact small samples can have on testing. The middle market is particularly challenging because disclosure of financial statements and loan performance is inconsistent. Our tests validated—to standards specified by the Basel II documents—that the new RiskCalc v3.1 features described in the previous sections of this paper did in fact produce higher performance in the RiskCalc models.

4.1   Model Power and Calibration

We consider validation in terms of two distinct dimensions of model quality. These dimensions are model power and model calibration. A model's power describes how well a model discriminates between defaulting ("bad") and non-defaulting ("good") borrowers. For example, if we were evaluating two models, A and B, each of which produced ratings of "good" and "bad," the more powerful model would experience a higher percentage of defaults (and a lower percentage of non-defaults) in its "bad" category. This model also would produce a higher percentage of non-defaults (and a lower percentage of defaults) in its "good" category. Thus, good rating systems (e.g., Moody's agency ratings) can have strong power to rank credit risk without necessarily being calibrated to default probabilities.

A model's calibration describes how well its predictions of default probability agree with actual outcomes. For example, it describes how close the model's default probability predictions match actual default rates, rather than describing how well the model differentiates defaulting borrowers from non-defaulting borrowers.

---

[24] Please see the RiskCalc v3.1 Modeling Methodology document for more detail.

4.2 Validation via Out-of-Sample Data

Moody's KMV uses a rigorous framework for model validation that emphasizes testing on data that was not included in the development sample. This data is referred to as out-of-sample data.

We structure out-of-sample data in a number of ways. For example, we developed an approach called walk-forward testing (Sobehart, Keenan and Stein, 2000; Stein, 2002), which allows users to test models and modeling methodologies while controlling for sample and time dependence. The technique reduces the chances that models will be "overfitted" since it never uses data in the testing that were used to fit model parameters. At the same time, this approach allows modelers to take greater advantage of the data by using as much of the information as possible to fit and to test the models. Other out-of-sample designs involve cross-validation and holdout samples. A more detailed discussion of validation approaches and experimental designs for testing models can be found in the Technical Document.

The following examples distinguish between results derived from in-sample and out-of-sample data. In addition to out-of-sample and out-of-time testing, we also employ other diagnostic techniques to manage the risk of overfitting. (An "overfitted" model performs very well in-sample on the data used to fit the model, but poorly out-of-sample on new data. Some causes of overfitting include incorporating too many ratios, capturing spurious relationships due to data problems and incorporating too much flexibility in modeling methods causing model instability.)

For example, basic econometrics suggests that using two explanatory variables that are highly correlated will yield large standard errors for parameter estimates of these variables. We account for these errors by measuring the variance inflation factors for each ratio included in our model in order to minimize multiple instances of collinearity in the model.[25]

Overfitting may also cause a model to test strongly in one setting (e.g., a specific industry or firm size classification) but poorly in other settings. We can test this scenario for overfitting by comparing the predictive power of the model to available alternative models across industry groups, firm size classifications, regions and time periods. For a further description of the results of these tests, please see the modeling methodology documentation for each model.

Table 5 provides a powerful example of RiskCalc v3.1 performance, in this case demonstrating the performance of the U.S. model *in sample* (Figures 6 & 7 and Table 6 present the out-of-sample results.). While for brevity we only show results for the U.S. model, we obtained similar results, to varying degrees, for our other models.

In Table 5, we see that with the credit cycle adjustment, the improvement in model performance is an increase of almost eight points of Accuracy Ratio at the one-year horizon compared with the previous version of RiskCalc. Relative to other available alternatives, the results were more dramatic. The new RiskCalc v3.1 model outperformed both the Private Firm Model and the Z-score model by double digits at both the one and five year horizons. Moreover, despite the power performance of RiskCalc v1.0 vis-à-vis other contemporary models, the Financial Statement Only (FSO) mode of the v3.1 model outperforms the old model by five points at both horizons.[26] Similarly, FSO outperforms RiskCalc v1.0 in terms of level validation as measured by the log-likelihood differences.

---

[25] Details of variance inflation factor (VIF) methodology can be found in RiskCalc Korea methodology document by Kocagil and Reyngold (2003).

[26] The corresponding accuracy ratios are 54.3% (FSO) vs. 49.5% (RCv1.0) for the one year horizon and 35.7% (FSO) vs. 30.7% (RCv1.0) for the 5 year horizon.

Moody's | K·M·V

TABLE 5   Power enhancements of the new RiskCalc v3.1 Model

	1-year Horizon		5-year Horizon	
	Accuracy Ratio	Lead in Log Likelihood*	Accuracy Ratio	Lead in Log Likelihood*
RiskCalc v3.1	57.0%	0	35.7%	0
RiskCalc v1.0	49.5%	403.6	30.7%	196.7
Private Firm Model	46.1%	517	23.2%	437.1
Z-score	42.3%	1356.5	24.7%	803.3

*Presents the increase in log likelihood of RiskCalc v3.1 over the alternative model. Larger values indicate that levels of RiskCalc v3.1 are better calibrated in relation to the alternative model.

Figure 5 presents the cumulative accuracy profiles for the one-year models corresponding to Table 5. These figures reveal that the power improvements are largely in the middle of the distribution relative to RiskCalc v1.0 (particularly for the one-year model). This result implies that both very good and very poor credits are correctly identified as such by both RiskCalc v1.0 and RiskCalc v3.1. The added discriminatory power is assessing the credit quality of credits that fall in the middle range.

FIGURE 5   Power of Alternative Models (one- and five-year) — U.S.
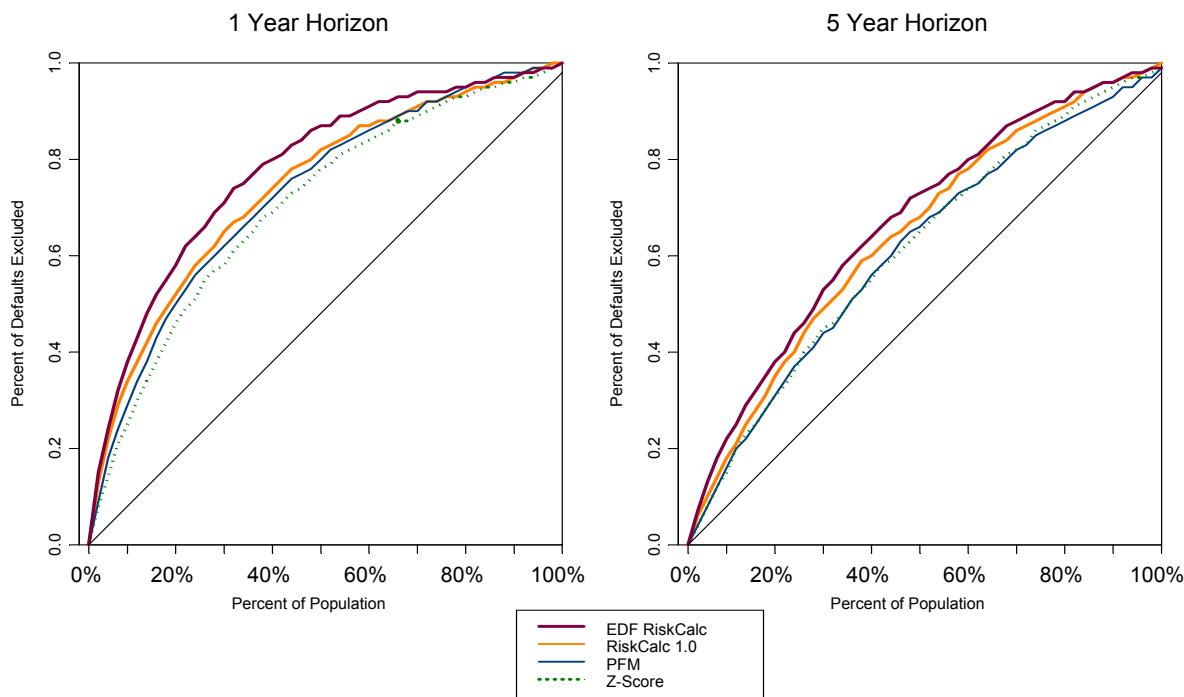
4.3   Testing Details

The first test we discuss is a K-fold analysis, which can be viewed as a test of model stability vis-à-vis different data segments. In this analysis, we divide firms into k sub-samples (we typically set k=5). Then we estimate the model on the sample excluding the observations in the set {k=1}. This model is used to score the observations in the set {k=1}. Such scores represent an *out-of-sample* model. We repeat this for each of the k sub-samples.

Afterwards, we combine the out-of-sample scores into one data set and calculate the accuracy ratio and the power curve. We then compare these results with the corresponding in-sample accuracy ratio and power curve. We also check to see whether the parameter estimates for each explanatory variable are stable across the different samples. Figure 6 presents the results of this analysis. Note that the model performance is maintained both in- and out-of-sample in the K-Fold analysis. The difference in AR between the in-sample and out-of-sample results is not bigger than one point in all cases. Further, RiskCalc v3.1 outperforms RiskCalc v1.0 in an out-of-sample context at both the one- and five-year horizons.[27]

FIGURE 6   Out-of-sample performance (one- and five-year) U.S. K-Fold
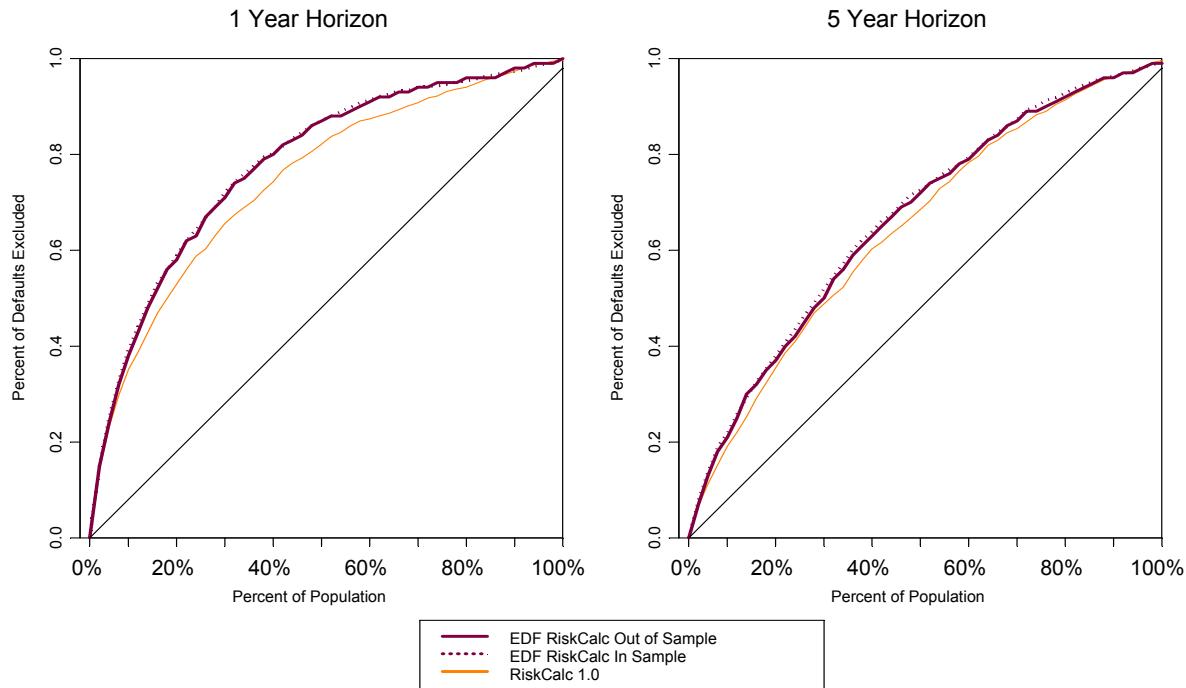
Note that the K-Fold testing does not control for time-dependence. Each of the *k* sub-samples contains data from all periods. As a result, if there were a particularly high period of default rates, this would be included in each of the *k* samples. Such testing does not give a true sense of the how the model would have performed during those volatile periods since the model is estimated with full information on those time periods.

An alternative out-of-sample test developed by Moody's KMV is a *walk-forward* analysis, which works in a similar fashion to the K-fold test, except that it controls for the effects of time. It proceeds as follows: We estimate the model up to a certain year and score the observations in the *next* year. These model scores are *out-of-time.* We then re-estimate the model including one more year of data and repeat the analysis for the next year and continue until the end of the sample. We combine these out-of-sample out-of-time scores into a single prediction set and calculate the accuracy ratio and the power curve for the combined set. We then compare with the corresponding in-sample accuracy ratio and power curve.

Note that no data from a future period is used in fitting the model and only data from future periods is used for testing it. As in the K-fold, we check to see if the parameter estimates are stable across the different samples. Figure 7 presents the results from this analysis. We observe that model performance is maintained. The difference in AR between the in-

---

[27] The out-of-sample ARs are 56.8% and 34.7% for the one-year and five-year models, respectively. These out-of-sample ARs are 0.7 and 1.0 points lower than the in-sample ARs and 6.5 and 3.6 points higher than RiskCalc v1.0, for the one and five year models respectively.

sample and out-of-sample results is no more than one point in all cases. Further, RiskCalc v3.1 outperforms RiskCalc v1.0 in an out-of-time context at both the one- and five-year horizons.[28]

FIGURE 7   Out-of-sample performance (one- and five-year) U.S. walk-forward
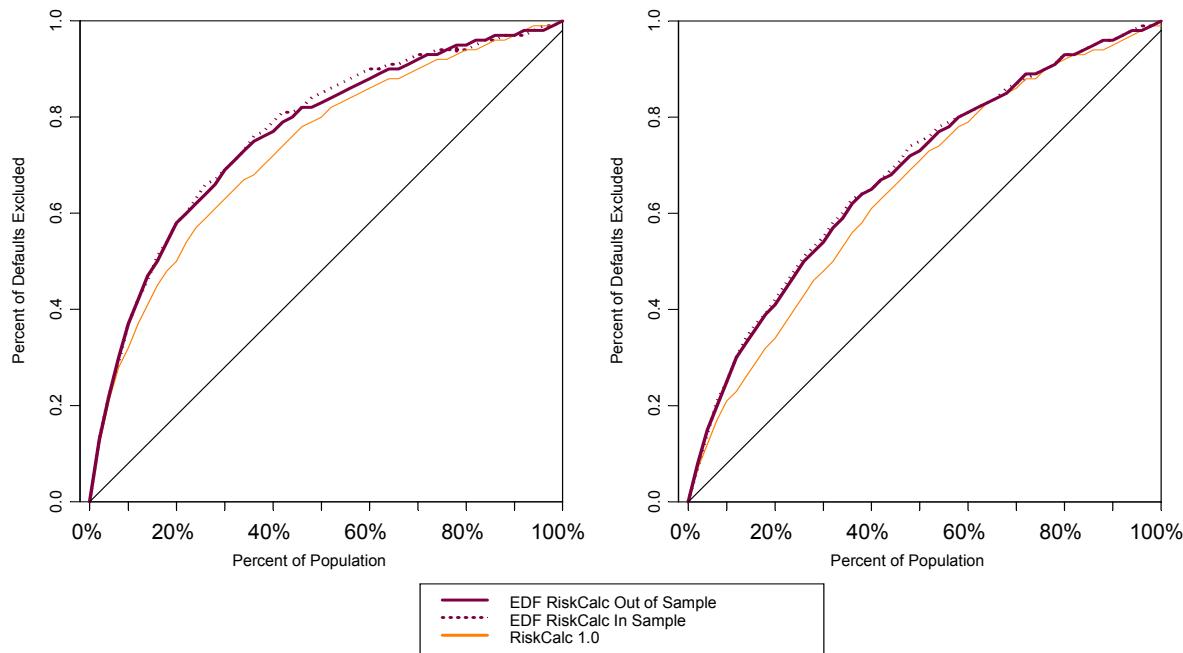
Finally, in one of our most rigorous tests to date, we report the results of a test in which we examine the power of the model on a dataset that became available only after the model was completed (i.e., a dataset that was not used for either developing or calibration of the model or for any of the other validation tests described here). This data set resulted from the most recent submission from contributors to the CRD and only became available in December of 2003, after the model was being implemented for production.

This newly available data has enabled us to conduct a purely out-of-sample test of model performance. The data included a new round of submissions from both new and existing CRD participants. We used this new data to construct a holdout sample, defined as firms that were never in the development or validation sample. This sample included more than 20,000 usable new observations and 500 usable new defaults.

Such a holdout sample allows for a pure out-of-sample testing for a number of reasons. For example, since the data was not available when RiskCalc v3.1 was developed, the holdout sample provides a test of the variable selection process and the model's functional form.[29]

---

[28] The out-of-sample ARs are 53.6% and 37.7% for the one-year and five-year models, respectively. These out-of-sample ARs are 1.0 and 0.6 points lower than the in-sample ARs and 5.9 and 6.0 points higher than RiskCalc v1.0, for the one and five year models, respectively.

[29] For a discussion of the issues surrounding exploratory analysis contaminating the holdout sample (also known as data-snooping), see Campbell et al. (1997, pages 523–524).

TABLE 6   Pure out-of-sample performance of the RiskCalc v3.1 Model

	One-year Horizon		Five-year Horizon	
	Accuracy Ratio	Lead in Log Likelihood*	Accuracy Ratio	Lead in Log Likelihood*
RiskCalc v3.1	60.8%	0	36.4%	0
RiskCalc v1.0	54.8%	193	28.4%	330
Z-score	43.3%	691	21.5%	862

*Presents the increase in log likelihood of RiskCalc v3.1 over the alternative model. Larger values indicate that levels of RiskCalc v3.1 are better calibrated in relation to the alternative model.

Table 6 reports the results of the analysis. Observe that RiskCalc v3.1 outperforms RiskCalc v1.0 by nearly 6.0 and 8.0 points in the 1-year and 5-year horizons, respectively. While not directly comparable because they are different samples, the gains in power are very consistent with those observed in Table 5.[30] Further, note that RiskCalc v1.0 continues to outperform the Z-score by a considerable margin. Finally, the likelihood gains of RiskCalc over every other model remain substantial as well. This provides perhaps the strongest evidence so far of the robustness and accuracy of RiskCalc v3.1.

4.4   Model Performance Over the Credit Cycle

Table 7 shows the relative performance of the model in each year for almost a decade. Both the RiskCalc v3.1 model and the Financial Statement Only model outperform all other models consistently across all years (with the exception of one tie in 1996). In addition, the RiskCalc v1.0 model substantially outperformed the available alternative models—Private Firm Model and Z-score—in 1999, 2000, and 2001 for the one-year and five-year horizons. These results confirm the out-of-sample performance of the RiskCalc modeling framework.[31]

TABLE 7   Model power over time: one-year horizon (AR = accuracy ratio)

	Percentage of defaults	AR RiskCalc v3.1	AR RiskCalc v1.0	AR Z-score	AR Private Firm Model
1993	1.2%	68.2%	64.5%	59.4%	65.1%
1994	3.3%	57.4%	55.4%	52.5%	43.6%
1995	5.6%	56.6%	53.0%	44.6%	45.5%
1996	7.1%	60.5%	60.1%	53.0%	50.9%
1997	11.3%	47.7%	44.2%	36.3%	37.0%
1998	20.8%	38.9%	35.0%	28.3%	32.1%
1999	23.6%	44.6%	39.6%	33.5%	35.1%
2000	19.8%	49.0%	43.6%	36.0%	38.8%
2001	7.2%	71.9%	66.5%	55.1%	63.0%

[30] The in-sample power gains are not directly comparable to the pure out-of-sample gains since they are based on different samples. See Stein (2002) for a discussion of issues in comparing validation results from tests on different samples.

[31] Recall that RiskCalc v1.0 was released in May 2000. As most financial statements come out at the end of the year, the majority of the defaults associated with 1999 financial statements are out-of-sample. Similarly, all of the defaults associated with 2000 and 2001 financial statements are out-of-sample.

Moody's | K·M·V

TABLE 8   Model power over time: five-year horizon (AR = accuracy ratio)

	Percentage of defaults	AR RiskCalc v3.1	AR RiskCalc v1.0	AR Z-score	AR Private Firm Model
1993	3.2%	45.1%	43.5%	33.6%	32.5%
1994	6.7%	44.7%	43.2%	34.9%	33.8%
1995	11.6%	38.3%	35.7%	26.2%	28.6%
1996	16.2%	36.1%	31.5%	22.0%	23.5%
1997	19.8%	33.6%	28.2%	22.2%	24.3%
1998	19.0%	37.4%	31.1%	26.2%	26.9%
1999	14.6%	42.3%	35.9%	31.7%	33.4%
2000	8.8%	48.5%	40.9%	34.8%	39.2%

5   ECONOMIC VALUE OF RISKCALC v3.1 MODEL POWER DIFFERENTIAL

Our benchmarks indicate that the new RiskCalc v3.1 model will deliver substantial economic value and profitability for users.

Our new models perform at a level that we anticipate will produce meaningful economic value for users and for the industry. Recent publications by Stein (2003) and Stein and Jordão (2003) provided a simplified framework for estimating an explicit dollar value based on the additional predictive power of a default model. While a discussion of these papers is beyond the scope of this article, key findings indicate:

- Powerful models are generally more profitable than weaker ones, regardless of the lending approach.

- A bank, by simply switching to a more powerful model (while keeping the same stream of borrowers) would enjoy meaningful benefits in terms of additional profit, even in the absence of competition.

- In a competitive environment, a bank using a more powerful model has an advantage over a bank using a weaker model because the latter would tend to suffer from adverse selection (gaming) towards customers that were correctly risk-priced by the bank with the powerful model.

- While banks using cutoff-based lending approaches (no origination for firms with EDF measures above "X") attained economically meaningful benefit from more powerful models, banks using even simplified risk-pricing approaches to lending enjoyed even greater benefits when competing against banks with less powerful models.

In order to demonstrate the magnitude of the impact of using the new RiskCalc v3.1 model, we performed a similar analysis to that of Stein and Jordão (2003), while making baseline assumptions about the lending environment. Below, Table 9 demonstrates the expected additional profit per dollar of credit granted that a user of the RiskCalc v3.1 model might enjoy relative to an alternative model.

TABLE 9   Estimated benefits (in basis points) of using the RiskCalc v3.1 model versus alternative models

	Switching (using cutoffs)	Competing using cutoffs	Competing using pricing
FSO mode of RiskCalc v3.1	*0.7 bps*	*1.4 bps*	*5 bps*
RiskCalc 1.0	*0.8 bps*	1.5 bps	18 bps
PFM	*1.6 bps*	3.1 bps	20 bps
Z-Score	*4.1 bps*	8.0 bps	25 bps

This table demonstrates the expected additional profit per dollar of credit granted that a user of the RiskCalc v3.1 model might enjoy relative to an alternative model. "Switching" implies that the banks prospective customers do not change but different firms may be granted/denied credit, because of results of the model. "Competing using cutoffs" means that the bank using the RiskCalc v3.1 model competes against a rival who uses the alternative with both banks lending based on cutoffs. (Note that in reality, the switching case is not a realistic one and could not be implemented in many markets since competitors are an unavoidable feature of those markets.) "Competing using prices" means that the bank using RiskCalc v3.1 model competes against a rival who uses the alternative with both banks lending based "on prices" derived from the models' predictions.[32]

The cells that are shown in bold are the most likely combinations available to users. Please note that these figures are meant to be estimates based on a specific set of (reasonable) assumptions, but that individual banks' lending environments, costs, and other variables differ and thus would be expected to change these results. One of the convenient features of the framework proposed by Stein (2003), however, is that the same model can be modified for use by banks with different cost and profit constraints.

To demonstrate the profit incentive for using our new model, we also applied the differences in profitability from Table 9 to a typical mid-sized bank from our middle-market Credit Research Database (CRD). To convert the average return analysis into a dollar value for a specific bank, we need to estimate the dollar value of new origination for the bank and to apply the average return estimate to this origination amount. Using the CRD, we were able to estimate the typical new issuance in 2002 for banks of various sizes. Such a bank made about $4.25 billion in loans in 2002 (new origination and renewal). Table 10 shows the economic value, in dollar terms, of adopting the RiskCalc v3.1 model. Again, the differences in profits are notable, particularly for banks using pricing approaches.

TABLE 10   Estimated benefits (in dollars of additional profit) of using the RiskCalc v3.1 model versus alternative models for a typical mid-sized bank ($4.24 billion in credit in 2002)*

	Switching (using cutoffs)	Competing using cutoffs	Competing using pricing
FSO mode of RiskCalc v3.1	*$ 297,500*	*$ 595,000*	*$ 2,125,000*
RiskCalc 1.0	*$ 340,000*	**$ 637,500**	**$ 7,650,000**
PFM	*$ 680,000*	**$ 1,317,500**	**$ 8,500,000**
Z-Score	*$ 1,742,500*	**$ 3,400,000**	**$ 10,625,000**

* Example client selected from Moody's KMV Credit Research Database

It is interesting to examine how banks would generate the additional revenue. For example, the bank using the RiskCalc v3.1 model to compete on pricing against a bank using Z-Score made the additional $10.6MM while lending to about 75 percent more borrowers (the more powerful model declines fewer non-defaulting firms), but incurring far fewer defaults (the more powerful model accepts fewer defaulting firms). *In other words, the RiskCalc v3.1 model bank's portfolio default rate was about one-fifth of that of its competitor.* In contrast, the bank using the RiskCalc v3.1 model for cutoff-based lending would have made about the same number of loans (96 percent) as a competitor using Z-Score for cutoff-based lending but would have experienced a default rate that was 25 percent lower, as they would have accepted and turned down different firms.

Since a large number of other factors may affect lending profitability beyond those discussed here, the profitability estimates from our simulation do not readily translate into good estimates of the realized profitability of the portfolio. However, by examining the differences between the estimates of profitability using each model, we can arrive at a reasonable first approximation to the dollar value that is attributable to the difference in model power*; this is the quantity we report. These estimates ignore the financial impact of the operational benefits of tracking more firms with the same analytic resources and implementing automated monthly monitoring to identify higher risk firms.

---

[32] Note that we are using the term "pricing" somewhat loosely here. In particular, the probabilities on which the "prices" are based are not risk-neutral probabilities but rather actuarial or "real world" probabilities. That said, a more complete translation from real-world to risk neutral probabilities can often be made with an estimate of the risk premium associated with a particular borrower (cf., Duffie and Singleton 2003).

6 SUMMARY AND CONCLUSIONS

RiskCalc v3.1 is the most powerful default prediction technology available for assessing middle-market credit risk.

Over the past decade, Moody's KMV has refined its techniques for gauging credit quality in the middle market. The result is the RiskCalc v3.1 model, the next-generation of private firm default-prediction technology.

Using extensive research and rich, proprietary data sets, we developed unprecedented insight into the drivers of default for private firms. RiskCalc v3.1 incorporates this learning by combining the RiskCalc v1.0 framework, the industry's leading middle-market modeling approach, with the Moody's KMV distance-to-default value, a proprietary measure that extracts forward-looking sector information from the equity markets. Our new default prediction tool equips you to convert equity information into credit signals. We also tested and introduced a number of additional innovations including a full continuous term structure of default rates and leading edge innovative approaches to validation, as well as features that regulators require under the New Basel Capital Accord.

The resulting RiskCalc v3.1 model is more intuitive to use and provides better indications of default probability than was possible in the past—and our testing confirms this. The RiskCalc v3.1 model outperforms all other models examined by substantial margins, both in terms of predictive power and in terms of the accuracy of the probabilities that are produced by the models. RiskCalc v3.1 will result in significant improvements in credit portfolio performance for users.

7 APPENDIX

Financial Statement Ratios used in RiskCalc v3.1 U.S., Canada, U.K., and Japan.

U.S.	Canada
PROFITABILITY • ROA • Change in ROA *LEVERAGE* • LTD to (LTD plus Net worth) • Retained Earning/ Current Liabilities *DEBT COVERAGE* • Cash Flow/Interest Expense *LIQUIDITY* • Cash and Marketable Securities/Assets *ACTIVITY* • Inventory/Sales • Change in AR Turnover • Current Liabilities/Sales *GROWTH* • Sales Growth *SIZE* • Total Assets	*PROFITABILITY* • ROA • Change in ROA *LEVERAGE* • LTD to (LTD plus Net worth) • Retained Earning to Current Liabilities *DEBT COVERAGE* • Cash Flow to Current Liabilities *LIQUIDITY* • Cash and Marketable Securities to Assets *ACTIVITY* • Inventory to Sales • Change in AR Turnover • Current Liabilities to Sales *GROWTH* • Sales Growth *SIZE* • Total Assets
U.K.	**Japan**
PROFITABILITY • Net P&L/ Turnover • Change in ROA *LEVERAGE* • Liabilities / Assets *DEBT COVERAGE* • Cash Flow/Interest Charges *LIQUIDITY* • Current Assets / Current Liabilities *ACTIVITY* • Trade Creditors to Turnover (Accounts Payable to Sales) • Change in Trade Debtors to Turnover (Change in Accounts Receivable to Sales) *GROWTH* • Turnover (Sales) Growth *SIZE* • Total Assets	*PROFITABILITY* • Gross Profit to Total Assets • Previous Year's Net Income to Previous Year's Net Sales *LEVERAGE* • Total Liabilities less Cash to Total Assets • Retained Earnings to Total Liabilities *LIQUIDITY* • Cash to Total Assets • Trade Receivables to Net Sales *ACTIVITY* • Inventory to Net Sales *INTEREST COVERAGE* • EBITDA to Interest Expense *Growth* • Sales Growth *SIZE* • Real Net Sales

Moody's | K·M·V

8 REFERENCES

1. The Basel Committee on Banking Supervision. *Third Consultative Paper* Bank for International Settlements, 2003 (http://www.bis.org/bcbs/bcbscp3.htm).

2. Bohn, Jeffrey T. "An Empirical Assessment of a Simple Contingent-Claims Model for the Valuation of Risky Debt." *The Journal of Risk Finance* 1, no. 4 (2000): 55-77.

3. Campbell, John Y., Andrew W. Lo and A Craig Mackinlay, *The Econometrics of Financial Markets,* Princeton University Press, 1997.

4. Crosbie, Peter J. and Jeff R. Bohn. "Modeling Default Risk." San Francisco: KMV, 2003.

5. Currie, Antony and Jennifer Morris. "And Now for Capital Structure Arbitrage." *Euromoney*, December 2002.

6. Duffie, Darrell and Kenneth J. Singleton. *Credit Risk: Pricing Measurement and Management*, *Princeton Series in Finance*. Princeton: Princeton University Press, 2003.

7. Dwyer, D.W and Roger M. Stein. "Technical Document on RiskCalc v3.1 Methodology." Moody's KMV, 2004.

8. Dwyer, Douglas and Roger Stein. "Inferring the Default Rate in a Population by Comparing Two Incomplete Default Databases." Moody's KMV, 2003.

9. Hamilton, David. "Default and Recovery Rates of Corporate Bond Issuers a Statistical Review of Moody's Ratings Performance 1970-2002." 52. New York: Moody's Investors Service, 2003.

10. Hastie, Trevor J., Rob J. Tibshirani. *Generalized Additive Models*. Vol. 43, *Monographs on Statistics and Applied Probability*. London: Chapman & Hall/CRC, 1990.

11. Hausman, J.A., Jason Abrevaya and F.M. Scott-Morton. "Misclassification of the Dependent Variable in a Discrete-Response Setting." *Journal of Econometrics* 87, no. 2 (1998): 239-70.

12. Keenan, Sean "Predicting Default Rates: A Forecasting Model for Moody's Issuer-Based Default Rates." In *Special Comment*, 20. New York: Moody's Investors Service, 1999.

13. Klein, J.P. and Moeschberger, M.L. *Survival Analysis: techniques for Censored and Truncated Data*, Springer Verlag, New York, 1997.

14. Kocagil, Ahmet and Alexander Reyngold. "Moody's RiskCalc™ for Private Companies: Korea." In *Rating Methodology*. New York: Moody's Investor Services, 2003.

15. Kurbat, Matthew and Irina Korablev. "Methodology for Testing the Level of the EDF™ Credit Measure." San Francisco: Moody's KMV, 2002.

16. Leland, Hayne E. and Klaus Bjerre Toft. "Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads." *Journal of Finance* 51, no. 3 (1996).

17. Longstaff, Francis A., and Eduardo Schwartz. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt." *Journal of Finance* 50, no. 3 (1995).

18. Merton, Robert C. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates." *Journal of Finance* 29 (1974): 449-70.

19. Nigrini, Mark J. "A Taxpayer Compliance Application of Benford's Law." *The Journal of the American Taxation Association* 18, no. 1 (1996): 72-91.

20. Nigrini, Mark J. *Digital Analysis Using Benford's Law*. Vancouver, BC: Global Audit Publications, 2000.

21. Pagan, Adrain and Aman Ullah. *Nonparametric Econometrics*, *Themes in Modern Econometrics*: Cambridge University Press, 1999.

22. Sobehart, Jorge R., Sean C. Keenan, and Roger M. Stein, 2000, Benchmarking Quantitative Default Risk Models: A Validation Methodology, (Moody's Investors Service).

23. Stein, Roger M., Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation, Moody's KMV, New York, 2002.

24. Stein, Roger M. "Power, Profitability and Prices: Why Powerful Models Increase Profits and How to Define a Lending Cutoff If You Must." Moody's KMV, November 2003.

25. Stein, Roger M., Ahmet E. Kocagil, Jeff Bohn, and Jalal Akhavein. "Systematic and Idiosyncratic Risk in Middle-Market Default Prediction: A Study of the Performance of the RiskCalc™ and PFM™ Models." Moody's Investor Services, February 2003.

26. Stein, Roger M., and Felipe Jordão. "What Is a More Powerful Model Worth?" Moody's KMV, August 11, 2003.

27. Stein, Roger. "Are the Probabilities Right? A First Approximation to the Lower Bound on the Number of Observations Required to Test for Default Rate Accuracy." New York: Moody's KMV, 2003.

28. Zeger, Scott L., and Kung-Yee Liang. "Longitudinal Data Analysis for Discrete and Continuous Outcomes." *Biometrics* 42 (1986): 121-30.

29. Zhou, Chunsheng. "A Jump-Diffusion Approach to Modeling Credit Risk and Valuing Defaultable Securities." Federal Reserve Board, 1997.

Moody's | K·M·V