MAY 13, 2005

MOODY'S KMV RISKCALC® V3.1 SOUTH AFRICA

MODELING**METHODOLOGY**

ABSTRACT

AUTHORS

Douglas W. Dwyer

Guang Guo

Moody's KMV RiskCalc® is the Moody's KMV model for predicting private company defaults. It covers over 80% of the world's GDP, has more than 20 geographic specific models, and is used by hundreds of institutions worldwide. While using the same underlying framework, each model reflects the domestic lending, regulator, and accounting practices of its specific region.

In January 2004, Moody's KMV introduced its newest RiskCalc modeling framework, Moody's KMV RiskCalc® v3.1. By incorporating both market (systematic) and company specific (idiosyncratic) risk factors, RiskCalc v3.1 is in the forefront of modeling middle-market default risk.

This document outlines the underlying research, model characteristics, data, and validation results for the RiskCalc v3.1 South Africa model.

Published by:
Moody's KMV Company

To contact Moody's KMV, visit us online at www.moodyskmv.com. You can also contact Moody's KMV through e-mail at info@mkmv.com, or call us by using the following phone numbers:

NORTH AND SOUTH AMERICA, NEW ZEALAND, AND AUSTRALIA:
1 866 321 MKMV (6568) or 415 874 6000

EUROPE, THE MIDDLE EAST, AFRICA, AND INDIA:
44 20 7280 8300

ASIA-PACIFIC:
852 3551 3000

JAPAN:
81 3 5408 4250

TABLE OF CONTENTS

1 INTRODUCTION

The Moody's KMV RiskCalc® v3.1 South Africa model is built using the results of extensive Moody's KMV research, including:

- Moody's KMV RiskCalc® v1.0 and the Moody's KMV Private Firm Model®
- Moody's KMV Credit Research Database™ (CRD, the world's largest and cleanest private company default database)
- Industry sector information, market information, and industry-specific default rates

RiskCalc v3.1 incorporates the structural and market-based comparables approach (used in the MKMV Private Firm Model), and the localized financial statement-based approach (used in RiskCalc v1.0). This allows RiskCalc v3.1 to blend market-based (systematic) information with detailed firm-specific financial statement (idiosyncratic) information to enhance the model's predictive power.

RiskCalc Modes

RiskCalc v3.1 allows a user to assess the risk of a private firm in two ways: Financial Statement Only (FSO) and Credit Cycle Adjusted (CCA).

The Financial Statement Only (FSO) mode delivers a firm's default risk based only on financial statements and sector information, adjusted to reflect differences in credit risk across industries. In this mode, the risk assessments produced by the model are relatively stable over time.

The Credit Cycle Adjusted (CCA) mode adjusts the default risk by taking into account the current stage of the credit cycle. The CCA adjustment is a sector-specific factor derived directly from the Moody's KMV public firm model's distance-to-default. The CCA model reflects the market's current assessment of the credit cycle and is a forward-looking indicator of default.

The CCA adjustment is specific to the firm's sector and country and is updated monthly. The CCA mode also has the ability to stress test EDF credit measures under different credit cycle scenarios – a requirement under Basel II.

2 DATA DESCRIPTION

The source of the data for RiskCalc v3.1 South Africa is Moody's KMV Credit Research Database™ (CRD). Moody's KMV collects data from participating institutions, working closely with them to understand the strengths and weaknesses of the data. As of April, 2005 the CRD contained 12.3 million financial statements, over 2.6 million unique private firms, and more than 175,000 default events worldwide. Moody's KMV uses this data for model development and testing purposes.

2.1 Definition of Default

RiskCalc® provides assistance to institutions and investors in determining the risk of default, missed payment, or other credit events. The process of developing the Basel Capital Accord stimulated debate around what constitutes an appropriate definition of default. RiskCalc applies the criteria used by most of the advanced economies in the world.

Default is defined as any of the following events:

- 90 days past due
- Bankruptcy and Liquidation
- Placement on internal non-accrual list
- Write-downs (e.g., unlikely to pay, charged-off loans, facilities called, special provisions, and restructured debt)

The most common types of default in the South African CRD are 90 days past due (31%), Bankruptcy & Liquidation (21%) and Charge-offs (18%).[1] We view this working definition to be consistent with the definition of default intended by the Basel Accord (see Paragraph 452 of "A Revised Framework").

2.2 Data Exclusions

Excluded Companies

The goal of the RiskCalc model is to provide an Expected Default Frequency™ (EDF) for private South African companies in the middle market. The firms and industries covered in the model must have similar default characteristics. To create the most powerful model for South African middle-market companies, companies that did not reflect the typical company in this market were eliminated. The following types of companies are not included in the data:

- **Small companies** – For companies with net sales of less than R250,000 (adjusted to account for inflation using a 2000 base year), future success is often linked to the finances of the key individuals. For this reason, they are not reflective of typical middle-market companies and are excluded from the database.
- **Financial institutions** – The balance sheets of financial institutions (banks, insurance companies, and investment companies) exhibit higher leverage than the typical private firm. The regulation and capital requirements of these institutions make them dissimilar to the typical middle-market company. Therefore, they are excluded from the database.
- **Real estate development companies** – The annual accounts of real estate development and investment companies provide only a partial description of the dynamics of these firms and, therefore, their likelihood of default. This is because their financial health often hinges on a particular development.[2]
- **Public sector and non-profit institutions** – Government run companies' default risks are influenced by the states' or municipalities' unwillingness to allow them to fail. As a result, their financial results are not comparable to other private firms. Not-for-profit financial ratios are very different from for-profit firms, particularly with regard to variables relating to net income.

[1] We classify defaults according to the most severe type of default. For example, a firm that first goes 90 days past due and then goes into bankruptcy is classified as a bankruptcy.
[2] There are many types of "project finance" firms whose success depends largely on the outcome of a particular project. We recommend using separate models for such firms. This characteristic is explicitly recognized in the Basel Capital Accord.

Excluded Financial Statements

The financial statements of smaller companies can be less accurate and of lower quality than those of larger companies. The financial statements in the CRD are cleaned to eliminate highly suspect financial statements. Plausibility checks of financial statements are conducted, such as assets not equal liabilities plus net worth, and financial statements covering a period of less than twelve months. If errors are detected, those statements are excluded from the analysis.

2.3 Descriptive Statistics of the Data

Overview of the Data

In order to build RiskCalc South Africa, we partnered with several leading South African financial institutions that provided financial statement data as well as default information. Figure 1 presents the distribution of South African financial statements and defaults by year in the CRD. A characteristic of this data set is that the majority of defaults are in 2002 and 2003. This characteristic is the result of recent increases in resources devoted to collecting default data, as opposed to being indicative of an economic downturn. During 2002 and 2003, South Africa was actually in an expansionary period. Table 1 summarizes the data used in the development, validation, and calibration of the RiskCalc v3.1 South Africa model.

FIGURE 1 Date Distribution of South African Financial Statements and Default Data

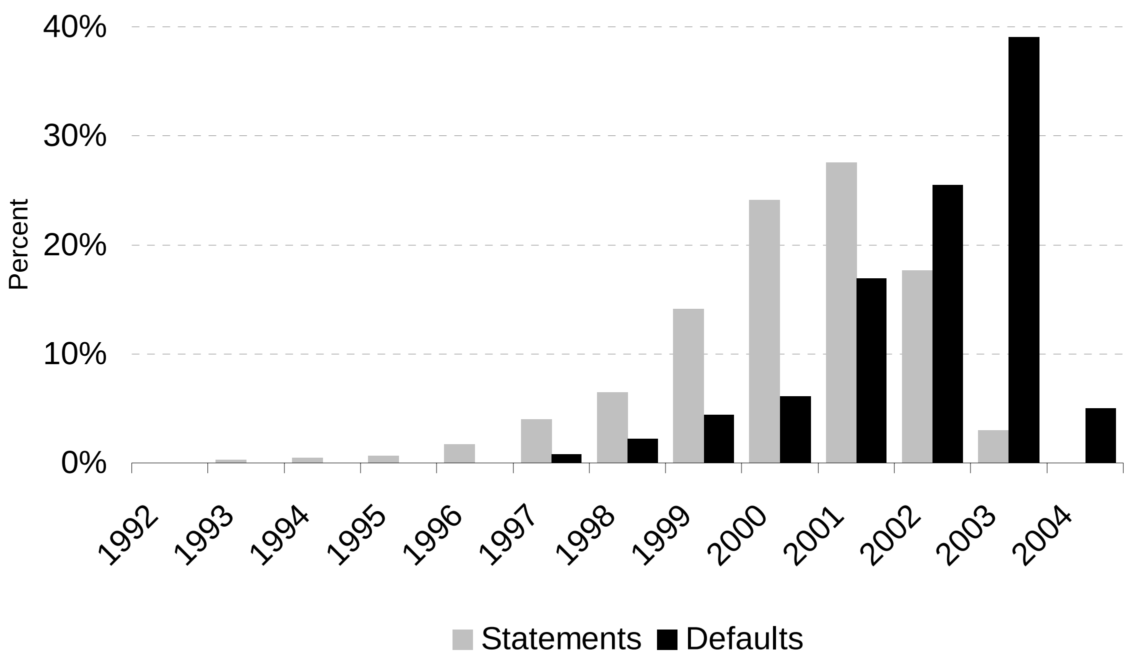


TABLE 1 Information on South Africa Private Firm Sample Data

CRD Private Firms	RiskCalc v3.1 South Africa	RiskCalc v1.0 U.S. Banks	RiskCalc v1.0 Portugal
Financial statements	52,000+	161,000+	69,000+
Unique number of firms	19,000+	17,000+	18,000+
Defaults	415	413	416
Time period	1992-2004	1986-1999	1993-2000

Robustness of the Data

In building a model, potential database weaknesses need to be examined. Not only does the database need to cover a large number of firms and defaults, but the defaults also need to be distributed among industries and company types covered. For example, if the database has significant numbers of small firms or firms in one particular industry and there are not sufficient defaults in those areas, the model may not be a good default predictor. The CRD used in developing the RiskCalc models has addressed both of these issues.

Figure 2 presents the distributions of South African defaults and firms by industry and the proportion of defaults in each industry. Figure 3 and Figure 4 present similar distributions by the size of firms measured as real total assets and real net sales in 2000 Rand, respectively. These figures demonstrate that the proportions of defaults in size groups or industry groups are roughly comparable to the proportion of firms in these groupings. The notable exceptions are Agriculture, which has a disproportionately fewer defaults, and Mining, Transportation, Utilities and Natural Resources which has disproportionately more defaults. The size distribution shows that the majority of firms (79%) hold greater than 1 million in assets. The proportion of firms with greater than 2.5 million in sales is 69% (where firms with less than R250,000 in real assets are excluded from the sample).

FIGURE 2 Distribution of South African Defaults and Firms by Industry

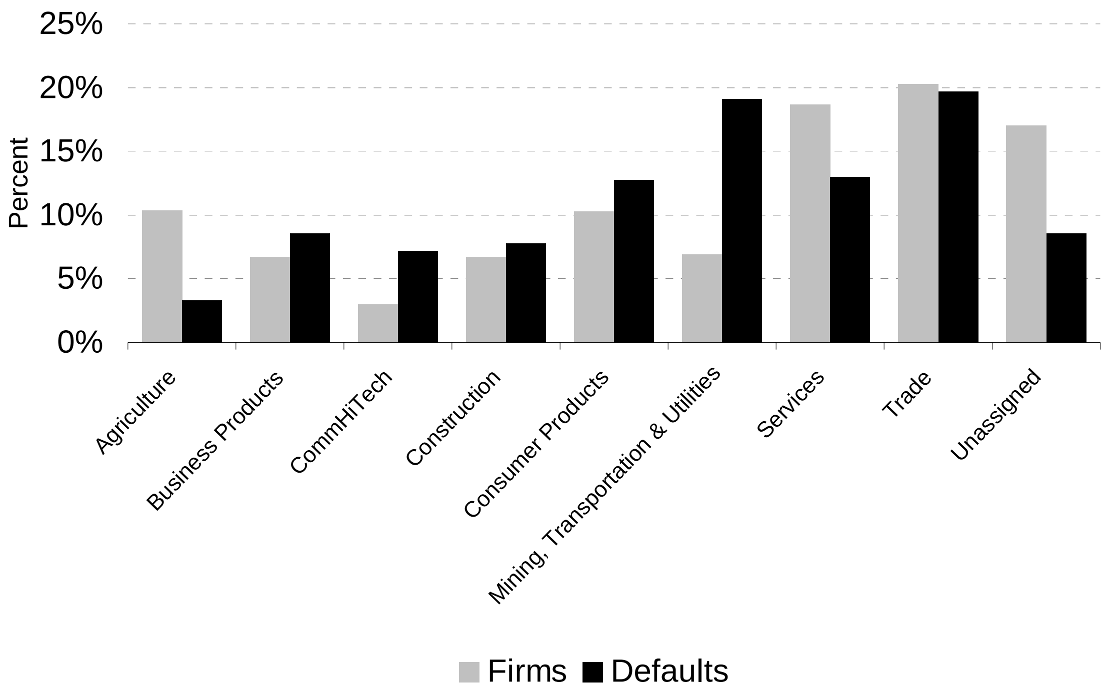


FIGURE 3 Size (as Total Assets) Distribution of Defaults and Firms

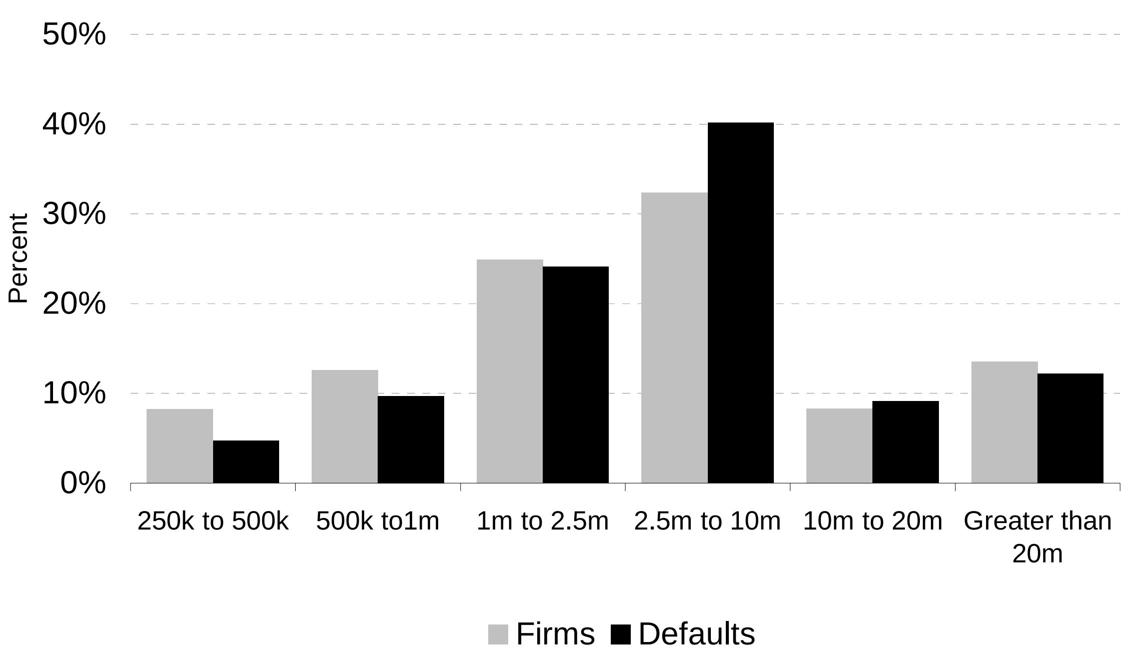
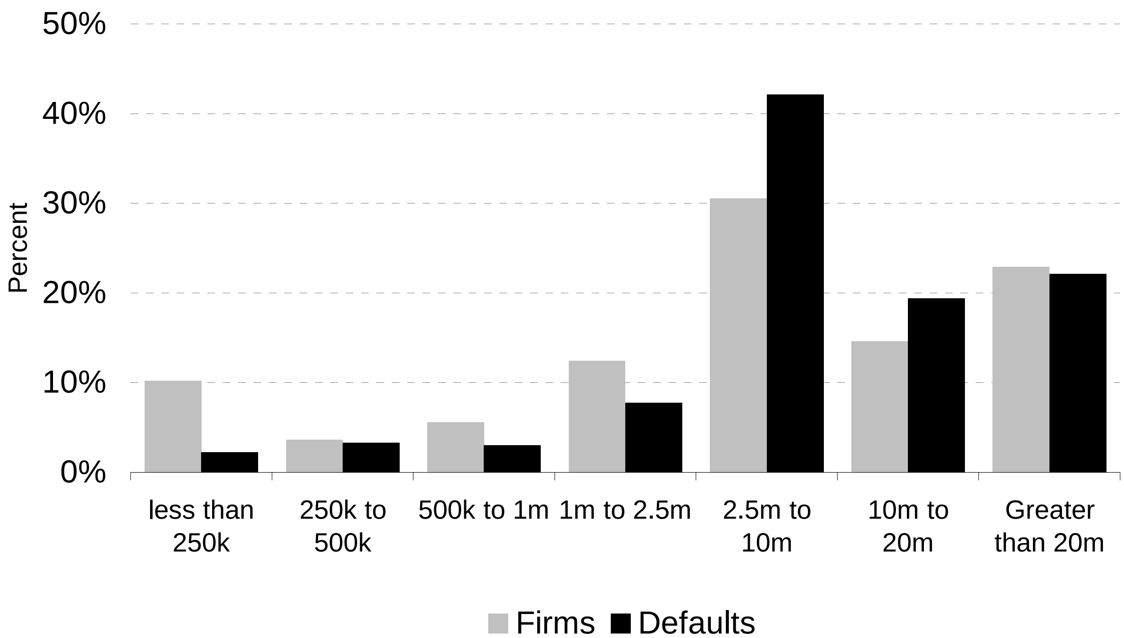


FIGURE 4 Size (as Net Sales) Distribution of Defaults and Firms



2.4 Cleaning the Data

In the development of a RiskCalc model, the first step is the collection of a large and appropriate database. In addition, data needs to be "cleaned" so that it is representative of the actual risk of the firms covered. MKMV has developed techniques for cleaning the database to improve the model results.

2.5 Central Default Tendency

Since most companies do not default, defaulting companies are relatively rare and thus more valuable in building a default prediction model. Much of the lack in default data is due to the data storage issues within financial institutions, such as defaulting companies being purged from the system after troubles begin, not all defaults being captured, or other sample errors. Also, if the date of default is uncertain, the financial statement associated with the firm may be excluded from model development, depending on the severity of the problem. These issues can result in a sample that has lower default rates than occur in the general population. If the underlying sample is not representative, then it needs to be adjusted for the true central default tendency.

In order to calculate the overall population default rate, we analyze information from both private and public records. The central default tendency is typically triangulated using two different approaches:

- Analysis of bank charge-offs and provisions.
- Reference to default rate and non-performing loans.

By estimating the central default rate from a variety of sources, the central tendency estimate is more accurate than that inferred directly from the development sample.

Bank Charge-off and Provisions

In order to determine the central tendency used in RiskCalc, MKMV analyzed public provisions data of major banks in South Africa. Banks make provisions for bad loans that are estimates of their expected write-offs. From the volume of losses and the volume of loans, an average default rate can be inferred given the loss given default (LGD):

$$\textit{Volume of Losses} = \textit{Volume of Loans} \times \textit{Probability of Default} \times \textit{LGD}$$

therefore

$$\textit{Probability of Default} = \textit{Volume of Losses} / (\textit{Volume of Loans} \times \textit{LGD}).$$

The analysis indicated an implied average yearly default rate of approximately 2.8% for South Africa assuming an LGD of 50%.[3]

Default Rate and Bank NPL.

An alternative data source is historical default rate data and Non-Performing Loan (NPL) ratios provided by partner financial institutions in South Africa. We found that the default rate was approximately 2.8%. This additional analysis confirmed the reasonableness of the chosen central default tendency.

Accordingly, in calibrating RiskCalc v3.1 for South African private companies, a central tendency of 2.8% was used for the one-year model.

Calculating a 5-year Central Default Tendency

There is a lack of publicly available data for direct calculation of the central tendency rate of a cumulative 5-year default probability. Based on extensive MKMV research, a 5-year cumulative default tendency is derived from the 1-year estimate. This research, combined with the information provided by the CRD, shows that the 5-year cumulative default rate is, on average, 4 times the level of the 1-year default rate. Therefore, 11.2% is used as the central default tendency for the 5-year model.

Central Default Tendency in FSO and CCA Modes

In the Financial Statement Only mode, the central default tendency remains constant over time. In Credit Cycle Adjusted mode, the central default tendency is equal to the central default tendency of the FSO mode when the effects of the credit cycle are neutral. When the forward-looking prediction of the credit cycle indicates increasing default risk, the central default tendency of the CCA mode will be larger, and when the effects of the credit cycle indicate reducing default risk, the central default tendency will be smaller.

[3] The "Revised Framework" uses 45% for senior claims and 75% for subordinated claims (paragraph 287 & 288) in the foundation approach.

3 MODEL COMPONENTS

The RiskCalc v3.1 model incorporates various components to determine the EDF credit measure. The inputs to the model include selection of the financial ratios and transforms of those ratios, the inclusion of industry information, and the credit cycle adjustment.

The development of a RiskCalc model involves the following steps:

1. Choosing a limited number of financial statement variables for the model from a list of possible variables.[4]
2. Transforming the variables into interim probabilities of default using non-parametric techniques.
3. Estimating the weightings of the financial statement variables, using a probit model, combined with industry variables.
4. Creating a (non-parametric) final transform that converts the probit model score into an actual EDF credit measure.

In FSO mode, the 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 $x_1,\ldots,x_N$ are the input ratios; $I_1,\ldots,I_K$ are indicator variables for each of the industry classifications; β and γ are estimated coefficients; Φ is the cumulative normal distribution; F and $T_1,\ldots,T_N$ are non-parametric transforms; and FSO EDF is the financial-statement-only EDF credit measure.[5] The Ts are the transforms of each financial statement variable, which capture the non-linear impacts of financial ratios on the default likelihood. (This is shown in Figure 5 and discussed in detail later in the document.) F is the final transform (i.e., the final mapping). The final transform captures the empirical relationship between the probit model score and actual default probabilities. We describe the final transform as calibrating the model score to an actual EDF credit measure. The difference between the FSO EDF and the credit cycle adjusted EDF is that in CCA mode the final transform is adjusted to reflect our assessment of the current stage of the credit cycle, while in FSO mode it remains constant.

3.1 Financial Statement Variables

Selecting the Variables

Our variable selection process starts with a long list of possible financial statement variables. The working list of ratios is divided into groups that represent different underlying concepts regarding a firm's financial status (see Table 3). A model is then built with at least one variable per group. When it is possible to increase model performance while maintaining model robustness, the model will use several variables from each group. Criteria that must be met for variables to be included in the final model are:

- Is the variable readily available?
- Are the definitions of the inputs to the variable unambiguous?

---

[4] These variables are often ratios but are not always ratios. For example, one measure of profitability is Net Income to Total Assets, which is a ratio, and one measure of size is Inflation Adjusted Total Assets, which is not a ratio.

[5] By non-parametric, we mean that the $T(x_i)$ 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.

- Is the meaning of the variable intuitive?
- Does the variable predict default activity?
- Is the variable not highly correlated with other variables in the model?

TABLE 2 Financial Statement Variables used in RiskCalc v3.1 South Africa

Category	Variable
Profitability	Net Income to Total Assets (ROA)
	Previous Year Net Income to Previous Year Total Assets (Previous Year ROA)
Leverage	[Current Liabilities + Long Term Debt] to Tangible Assets
Debt Coverage	Cash Flow* to Financial Charges
Growth	Sales Growth
Liquidity	Cash** to Total Assets
Activity	Inventories to Cost of Goods Sold
Size	Total Assets

* Cash Flow is operating cash flow and is implemented as EBITDA plus changes in accounts payable less changes in accounts receivable less changes in inventories.

**Cash includes marketable securities.

TABLE 3 Groupings of Financial Statement Ratios

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. → *High leverage increases the probability of default.*

Debt Coverage is the ratio of cash flow to interest payments or some other measure of liabilities. → *High debt coverage reduces the probability of default.*

Growth variables are typically the change in ROA and sales 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 to sales. These ratios may 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 2000 South Africa Rand). → *Large firms default less often.*

Variable Transforms

Once the variables are selected, they are transformed into a preliminary EDF value. Figure 5 presents the transformations used in the model. The horizontal axis gives the percentile score of the ratio and the vertical axis gives the default probability of that ratio in isolation (univariate). The percentile score represents the percent of the database that had a ratio below that of the company (e.g., if liquidity is in the 90th percentile that means that 90% of the sample had lower liquidity than that firm).

The shape of the transformation indicates how significantly a change in level impacts the EDF value. If the slope of the transform is steep, a small change will have a larger impact on risk than if the slope were flat.

- For the **Profitability** group, two ratios are included, ROA and Previous Year ROA. The transforms for both ratios are downward sloping, and the slopes become flatter as profitability increases (Figure 5). Such a transform indicates that more profitable firms have lower default probabilities, but the impact diminishes as profitability increases.

- For the **Leverage** group, the transform for [Current Liabilities + Long Term Debt] to Tangible Assets is upward sloping, as higher leverage indicates increased credit risk. The slope becomes less positive after the ratio reaches median percentiles and regains its steepness as the ratio increases further. This nonlinear pattern implies that as firms' leverage ratios drop into the very low and very high percentiles, the change in percentile has a larger impact than when the ratio is in the median percentiles (Figure 5).

- The **Debt Coverage** variable is Cash Flow to Financial Charges. The variable is downward sloping, indicating that large values of debt coverage lower the probability of default. The slopes become flatter as debt coverage ratios decreases, indicating that the impact of debt coverage diminishes when debt coverage ratios are small (Figure 5).

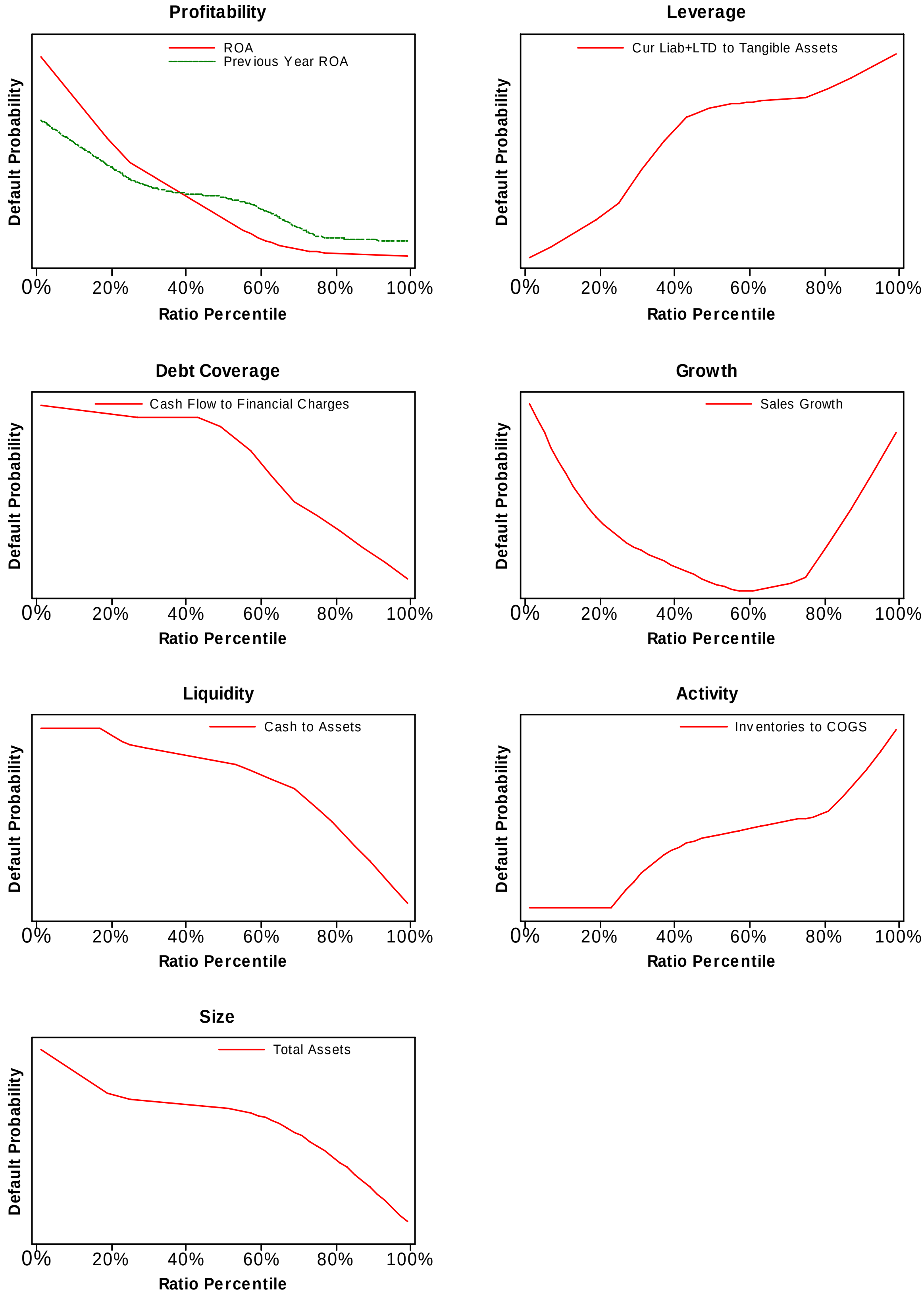
- The **Growth** variable is Sales Growth, which is "U shaped," indicating that large increases or decreases in sales are associated with higher default probabilities (Figure 5).

- For the **Liquidity** group, the transform for Cash to Total Assets is downward sloping with a flattened slope in very low percentiles, which means the more liquidity (higher cash to asset ratio) a firm has, the lower its default risk (Figure 5).

- For the **Activity** group, Inventories to Cost of Goods Sold is upward sloping indicating that high values of this ratio are associated with higher default probabilities (Figure 5). In addition, the slope is steepened in the high percentiles region (i.e., 80%), showing that the marginal impact of this ratio on credit risk jumps as the ratio grows very high.

- The **Size** variable is Total Assets. This variable's transformation is downward sloping (Figure 5). This indicates that larger firms have lower default probabilities.

FIGURE 5 Transformations of Financial Statement Variables Used in the Model



3.2 Model Weights

Importance

The relative value of each variable used in calculating an EDF credit measure is important in understanding a company's risk. The non-linear nature of the model makes the weight of the variables more difficult to determine since the actual impact on the risk depends on the coefficient, the transformation shape, and the percentile ranking of the company. The model weights, therefore, are calculated based on the average EDF value for the transformation and its standard deviation. A variable with a flat transformation could have a low weight, even if the coefficient is large (Figure 5).

Calculation of Weights

To calculate the weighting of a variable, the EDF credit measure for a theoretical firm with all its variables at the average transformation values is computed. The variables are then increased one at a time by one standard deviation. The EDF change for each variable (in absolute value) is computed and added together. The relative weight of each variable is then calculated as the EDF level changes for that variable as a percent of the total change in EDF level. This gives the variable with the biggest impact on the EDF level the biggest weight, and the variable that has the smallest impact on the EDF level the smallest weight. Since the weights are a percentage of the total EDF value, they sum to 100%. The weight of each category is the sum of the weights of each variable in that category. Table 4 presents the weights in RiskCalc v3.1 South Africa.

TABLE 4 Risk Drivers in RiskCalc v3.1 South Africa

RiskCalc v3.1 South Africa	
Risk Drivers	**Weight**
Profitability ROA Previous Year ROA	24.50%
Liquidity Cash to Total Assets	18.60%
Leverage [Current Liabilities + Long Term Debt] to Tangible Assets	17.58%
Growth Sales Growth	12.85%
Activity Inventories to Cost of Goods Sold	9.96%
Size Total Assets	8.63%
Debt Coverage Cash Flow to Financial Charges	7.89%

3.3 Industry Adjustments

While the variables included in the RiskCalc model explain most of the risk factors, the relative importance of the variables can be different among industries. In the FSO mode of RiskCalc v3.1 South Africa, the EDF value is adjusted for industry effects. In order to determine the industry adjustment, we first included indicator variables for each industry with the exception of Trade, which was used as the reference group. We chose Trade as the reference group because it is the largest sector in the economy. In the initial specification, the indicator variables for Business Products, Consumer Products, and Services were found to be statistically indiscernible from Trade. Therefore, the indicator variables for these sectors were omitted from the final specification. After removing these indicator variables, we found that Agricultural firms had lower default probabilities while Construction, Communications & Hi-Tech, and Mining, Transportation, Utilities and Natural Resources firms have higher default probabilities after controlling for the other variables in the model.

Table 5 presents the increase in model power and accuracy from introducing industry controls into the FSO model. Both the power and the accuracy of the EDF credit measure increase, as measured by the accuracy ratio and the gain in log likelihood. The large gain in likelihood indicates that the industry controls are important in producing an accurate EDF value. Table 6 presents the average EDF value by industry for the development sample in June of 2003. The highest default probabilities are in Construction and lowest are in Business Products.

TABLE 5 Increase in Model Power and Accuracy from Introducing Industry Controls

	One-Year Model		Five-Year Model	
	Accuracy Ratio*	Relative increase in Log Likelihood**	Accuracy Ratio	Relative increase in Log Likelihood
FSO mode without industry controls	41.44%		32.00%	
FSO mode with industry controls	45.10%	77.52***	36.05%	184.24***

* In this table, and hereafter, Accuracy Ratio (or AR) is the measure of the model's ability to rank order credits.

** Increases in log likelihood measure the extent to which the model's EDF values match observed default rates.

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

TABLE 6 Average EDF Credit Measure in 2003 June by Sector

Sector	Average 1-year EDF	Average 5-year EDF
Agriculture	2.16%	8.99%
Business Products	1.43%	7.25%
Communications and Hi Tech	2.15%	8.91%
Construction	3.14%	11.76%
Consumer Products	2.52%	10.70%
Mining, Transportation, Utilities and Natural Resources	2.95%	11.16%
Services	2.46%	9.53%
Trade	2.10%	9.27%

3.4 Credit Cycle Adjustment

EDF credit measures are impacted not only by the financials of a company, but also by the general credit cycle in the economy. To capture this effect, RiskCalc v3.1 South Africa includes a credit cycle adjustment factor. The credit cycle adjustment is designed to incorporate the current position of the credit cycle into the estimate of private firm default risk.

Selecting an Adjustment Factor

The RiskCalc v3.1 model uses the distance-to-default (DD) calculation from the Moody's KMV public firm model. This measure is specifically designed to be a forward-looking indicator of default risk. It extracts signals of default risk from the stock market performance of individual firms (cf. Crosbie and Bohn, 2003). This measure was chosen because it is available for a large universe of industries and it has been extensively validated.

If the distance-to-default for public firms in an industry indicates a level of risk above the historical average for that industry, then the private firms' EDF values in that industry are adjusted upward. Conversely, if the level of risk is below the historical average, then the private firms' EDF values are adjusted downward. When the credit cycle adjustment factor is neutral, the CCA EDF value coincides with the FSO EDF value.

Adjustment Factor Used in the Model

For the South African model, the distance-to-default (DD) factor for each industry is a weighted (by size) average index. The average is based on an aggregation of distance-to-default for all South African firms in each industry. In the event that a firm cannot be associated with a specific industry, the model uses a credit cycle adjustment based on an aggregation of all public South African firms.[6] Figure 6 provides evidence that the DD factor increases with insolvencies in South Africa.[7]

---

[6] In order to ensure a sufficient number of publicly traded South African companies we did the following: (i) treated business products, consumer products, and communications and hi-tech as one group; (ii) consolidated services and trade into one group; (iii) used an adjustment factor based on all South African public firms for agriculture and construction.

[7] The time-series on insolvencies represents the moving average of these series over the next 12 months. The time-series data of insolvencies of South Africa can be accessed via http://www.statssa.gov.za.

FIGURE 6 South Africa DD Factor and Insolvencies Data: 1993-2003

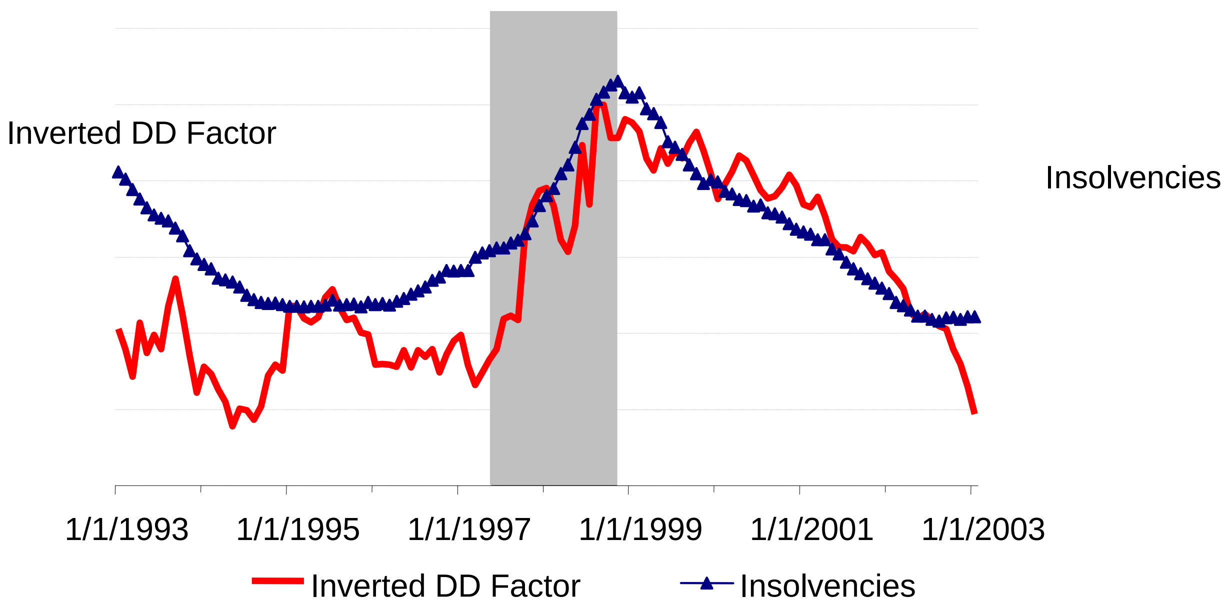


Figure 6 gives the DD factor (solid line) against the historical insolvencies data (triangle dotted line) for South Africa. The shaded area corresponds to a recessionary period between April 1997 and November 1998.[8]

4 VALIDATION RESULTS

Once a model is developed, it must be shown to be effective in predicting defaults. In this section, we present results on the model's ranking power and the accuracy of its predicted EDF credit measure (the model's ability to estimate the EDF level correctly.

The tests need to check not only the model's effectiveness, but also its robustness and how well it works on data outside the sample. To do out-of-sample testing, we performed walk-forward and k-fold analyses. The results of the testing show that the model is uniformly more powerful than other models across different time periods, sectors, and size classifications.

Table 7 presents the in-sample overall measures of power for RiskCalc v3.1 South Africa versus alternative models. Relative to available alternatives, the results were dramatic. The RiskCalc v3.1 model outperformed the Z-score model (Altman, Hartzell and Peck, 1995) by more than 15 percentage points at both the 1-year and 5-year horizons.[9] It outperforms the RiskCalc v1.0 North America Model by 12 and 6 percentage points at the 1-year and 5-year horizons, respectively.

[8] Source: Economic Cycle Research Institute (ECRI), New York City.

[9] The power results on the model run in CCA mode are skewed due to a data collection issue: the majority of defaults in the data are from 2001-2002, which represents a period of expansion in the South African economy. Consequently, the results presented here are for the model in FSO mode.

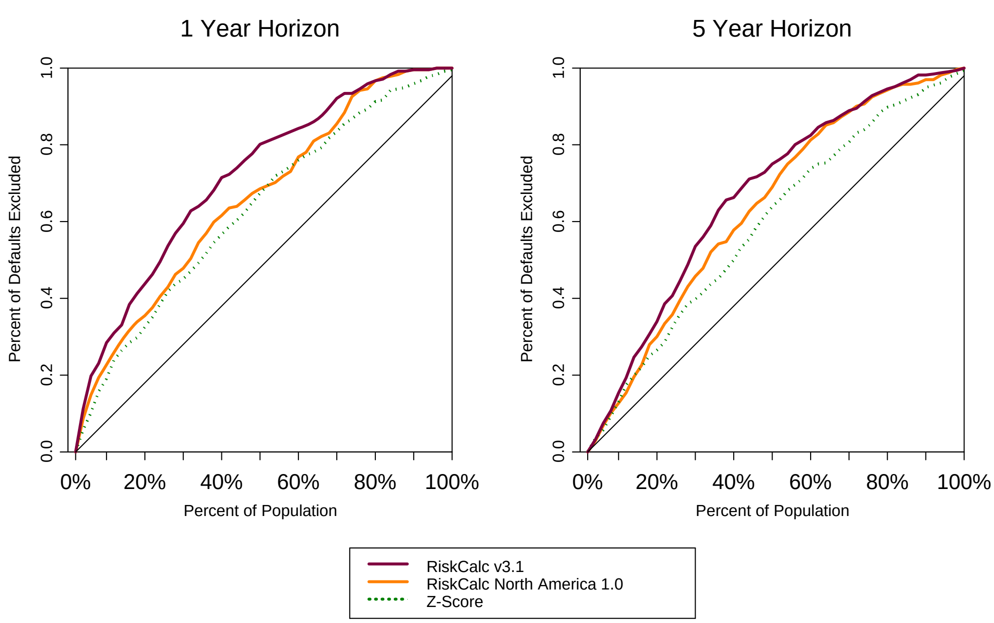
TABLE 7 Power Enhancements of the RiskCalc v3.1 South Africa Model

Model	Accuracy Ratio	
	One-year Model	Five-year Model
RiskCalc v3.1 Model	45.1%	36.1%
RiskCalc v1.0 Model North America	33.4%	29.7%
Z-score	27.7%	20.4%
ROA-Leverage*	23.8%	20.4%

* Falkenstein (2000, page 73) found the simple difference of ROA and Leverage to be an effective indicator of credit quality.

Figure 7 presents the cumulative accuracy profiles for the 1- and 5-year models corresponding to Table 7. The power improvements are significant in low and middle region relative to RiskCalc North America v1.0. This result implies that RiskCalc v3.1 South Africa is more powerful than alternative models.

FIGURE 7 Power of Alternative Models (1- and 5-year) — South Africa



4.1 Correlations and Variance Inflation Factors

To ensure model robustness, the model must be tested for excessive multicollinearity, which occurs if a number of the variables used in the model are highly correlated. Excessive multicollinearity can cause instability in parameter estimates. To check for this issue, the correlation coefficients (Table 8) for the financial statement ratios in the model and the variance inflation factors (Table 9) are computed on the transformed variables (see Figure 5).

Model Results

The highest correlation coefficient is between [Cash Flow to Financial Charges] and [ROA] (0.463). The next highest coefficient is between [Cash Flow to Financial Charges] and [Cash to Assets] (0.326). Such coefficients are well below what we would typically consider indications of multicollinearity and this finding is also verified by the VIF analysis.

TABLE 8 Correlations Among the Transformed Input Factors (Spearman Rank)

	Inventories/Cost of Goods	Cash Flow/Financial Charges	(Current Liabilities + Long Term Debt)/Assets	Sales Growth	Cash /Total Assets	Previous Year ROA	ROA	Total Assets
Inventories/Cost of Goods	1.000							
Cash Flow/Financial Charges	0.075	1.000						
(Current Liabilities + Long Term Debt)/Assets	(0.064)	0.048	1.000					
Sales Growth	(0.032)	0.254	(0.013)	1.000				
Cash / Total Assets	0.056	0.326	0.030	0.146	1.000			
Previous Year ROA	0.010	0.235	0.101	0.188	0.146	1.000		
ROA	0.026	0.463	0.133	0.225	0.186	0.284	1.000	
Total Assets	(0.278)	0.039	0.104	0.107	0.164	0.091	0.090	1.000

The Variance Inflation Factors (Table 9) for the financial statement variables represent how much of the variation in one independent variable can be explained by all the other independent variables in the model. The correlation coefficient, however, measures only the relationships between two variables. As shown in Table 9, the estimated VIF values in the South Africa model are notably below the threshold levels of 4 to 10 that are commonly used in VIF analysis when testing for presence of multicollinearity.[10] These findings indicate that the model variables do not present any substantial multicollinearity.

[10] As Woolridge (2000) shows, VIF is inversely related to the tolerance value (1-R^2), such that a VIF of 10 corresponds to a tolerance value of 0.10. Clearly, any threshold is somewhat arbitrary and depends on the sample size. Nevertheless, if any of the R^2 values are greater than 0.75 (so that VIF is greater than 4.0), we would typically suspect that multicollinearity could be a problem. If any of the R^2 values are greater than 0.90 (so that VIF is greater than 10) we then conclude that multicollinearity is likely to be a serious problem.

TABLE 9 Variance Inflation Factors

Variable	VIF
Cash Flow to Financial Charges	1.46
ROA	1.36
Cash / Total Assets	1.19
Total Assets	1.17
(Current Liabilities + Long Term Debt)/Tangible Assets	1.16
Inventories/Cost of Goods	1.14
Previous Year ROA	1.13
Sales Growth	1.05

4.2 Model Power by Industry and Size Groups

It is important to test the power of a model not only overall, but also among different industry segments and firm sizes. Table 10 and Table 11 present the power comparisons by sector for the 1-year and 5-year models, respectively. Results for industries that represent fewer than 10% of the defaults in the sample are not reported as such results are unreliable due to small sample issues.

At both the one and five year horizon (Table 10), RiskCalc v3.1 South Africa outperforms Z-score in all sectors. The highest power in the 1-year model is found in Mining, Transportation, Utilities and Natural Resources (53.5%) while the lowest is found in Consumer Products (29.7%). At the 5-year horizon (Table 11), the highest power is in Mining, Transportation, Utilities and Natural Resources (58.5%) and the lowest is in Trade (33.1%).

TABLE 10 Model Power by Industry: 1-year Model

	Percentage of Defaults	AR* RiskCalc v3.1	AR Z-score
Consumer Products	14.8%	29.7%	28.1%
Mining, Transportation, Utilities and Natural Resources	21.8%	53.5%	32.8%
Services	14.4%	43.9%	17.6%
Trade	21.3%	40.5%	27.7%

*AR = accuracy ratio

TABLE 11 Model Power by Industry: 5-year Model

	Percentage of Defaults	AR RiskCalc v3.1 South Africa	AR Z-score
Consumer Products	14.6%	34.4%	18.0%
Mining, Transportation, Utilities and Natural Resources	20.5%	48.5%	21.5%
Services	13.9%	36.2%	18.5%
Trade	21.9%	33.1%	17.8%

Table 12 and Table 13 present the power comparisons by firm size for the 1-year and 5-year models, respectively. RiskCalc v3.1 South Africa outperforms Z-score in all size groups. The highest power is among the largest firms which may be indicative of relatively high quality financial statements among such firms.

TABLE 12 Model Power by Size: 1-year Model

	Percentage of Defaults	AR RiskCalc v3.1 South Africa	AR Z-score
<to R1mm	13.0%	45.6%	35.6%
R 1mm to R2.5mm	23.3%	41.6%	15.9%
R2.5mm to R5mm	25.3%	43.1%	25.4%
R5mm to R15mm	22.1%	36.2%	22.3%
>R15mm	16.2%	64.3%	49.2%

TABLE 13 Model Power by Size: 5-year Model

	Percentage of Defaults	AR RiskCalc v3.1 South Africa	AR Z-score
<to R1mm	14.3%	34.7%	27.5%
R 1mm to R2.5mm	23.3%	32.5%	18.5%
R2.5mm to R5mm	27.6%	33.6%	18.0%
R5mm to R15mm	22.6%	39.6%	18.4%
>R15mm	12.2%	45.6%	33.4%

4.3 Power Performance over Time

Since models are implemented at various points in a business cycle by design, power tests by year (Table 14 and Table 15) were conducted to examine whether or not the model performance is excessively time dependent.

Table 14 and Table 15 present the results from this analysis at the 1- and 5-year horizons, respectively. The accuracy ratio of RiskCalc v3.1 South Africa is compared with Z-score for each year. As shown in these tables, throughout the whole period, RiskCalc v3.1 South Africa consistently outperforms Z-score by a considerable margin at both 1-year and 5-year horizon.

TABLE 14 Model Power over Time: 1-year Horizon

	Percentage of Defaults	AR* RiskCalc v3.1 South Africa	AR Z-score
1999	15.3%	67.4%	38.1%
2000	22.3%	39.8%	24.9%
2001	33.9%	40.0%	20.7%
2002	13.0%	37.7%	29.0%

*AR = accuracy ratio

TABLE 15 Model Power over Time: 5-year Horizon

	Percent of Defaults	AR RiskCalc v3.1 South Africa	AR Z-score
1997	11.1%	31.5%	16.7%
1998	17.2%	41.3%	24.6%
1999	21.2%	38.8%	20.6%
2000	20.6%	29.1%	13.1%
2001	15.2%	39.2%	22.1%

4.4 Out of Sample Testing: K-fold Tests

The model exhibits a good power in distinguishing good credits from bad ones (in Table 7), but whether this power is attributable to the overall model effectiveness or the impact of a particular sub-sample also needs to be tested. A standard test for evaluating this is the "*k*-fold test," which divides the defaulting and non-defaulting companies into *k* equally sized segments. The model is then run on *k-1* sub-samples and these parameter estimates are used to score the *k-th* sub-sample. We repeat this procedure for all possible combinations, and put the *k* scored "out-of-sample" sub-samples together to calculate an accuracy ratio on this combined data set.

Results

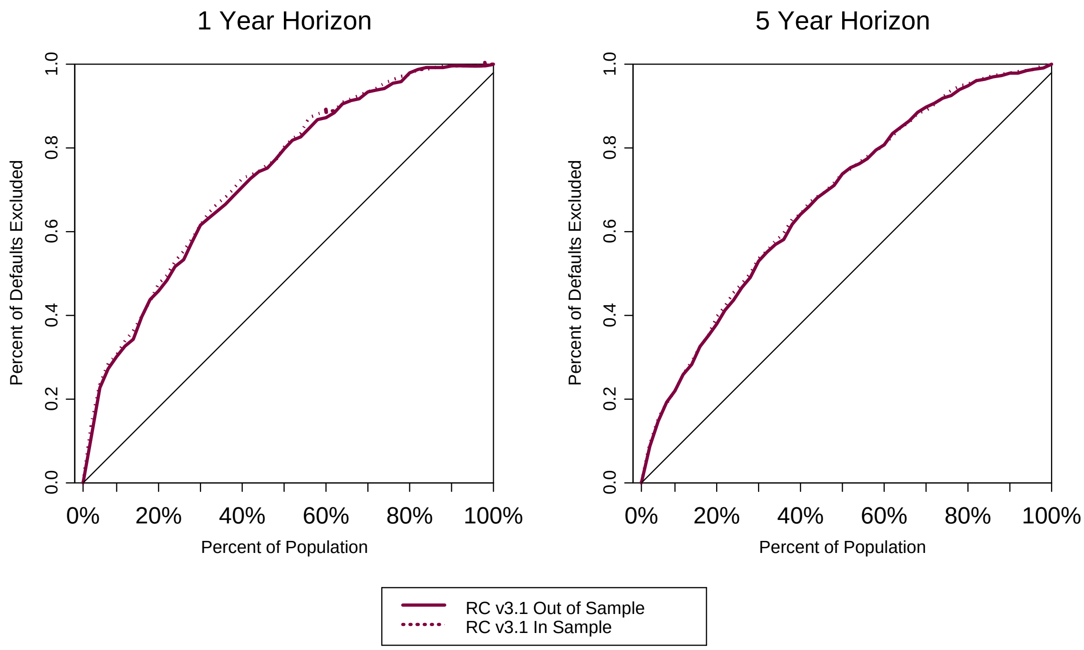
Table 16 summarizes the *k*-fold test results (with *k*=5). The reported figures are the accuracy ratios by the corresponding sample and time horizons. The out-of-sample model consistently out-performs Z-score. Figure 8 presents the cumulative accuracy profiles associated with the overall "out-of-sample" results against the in-sample results. The model performance is maintained both in- and out-of-sample in the *k*-fold analysis.

TABLE 16 RiskCalc v3.1 South Africa K-fold Test Results

	Out of Sample AR*		Z-Score	
	1-year AR	5-year AR	1-year AR	5-year AR
Subsample 1	44.6%	39.7%	31.5%	26.1%
Subsample 2	40.1%	35.9%	19.5%	19.8%
Subsample 3	40.7%	39.0%	30.4%	20.5%
Subsample 4	49.2%	42.6%	25.7%	15.5%
Subsample 5	37.3%	40.3%	32.5%	23.9%
K-fold Overall	44.2%	37.0%	27.7%	20.4%
In-sample AR	45.5%	36.1%		

*AR = accuracy ratio

FIGURE 8 RiskCalc v3.1 South Africa K-fold



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 because the model is estimated with full information on those time periods.

4.5 Walk-Forward Tests

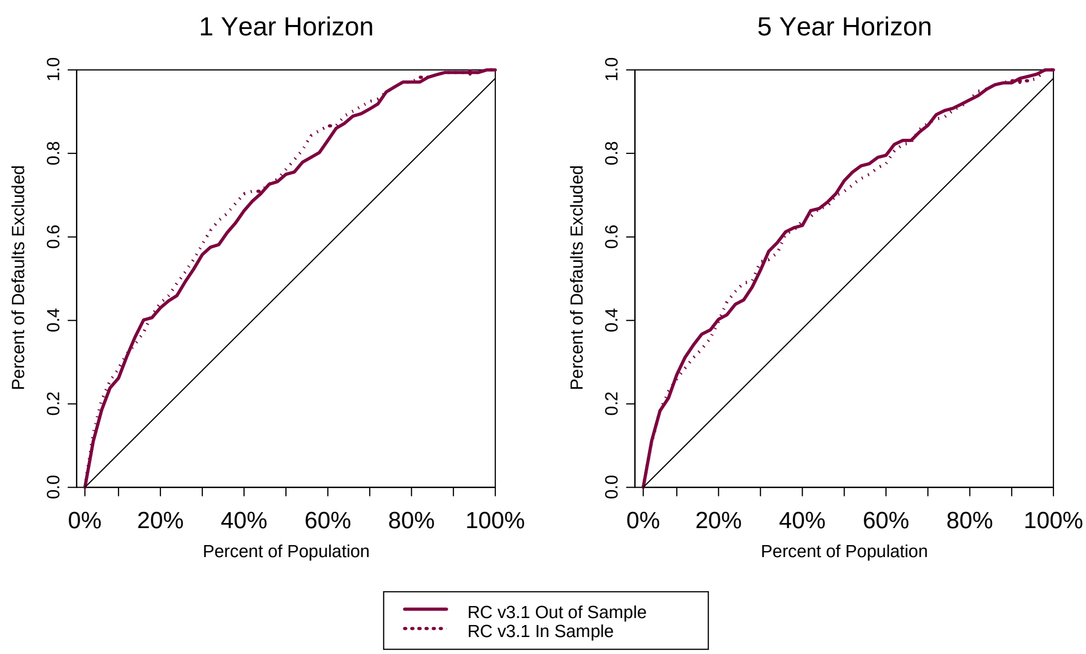
An alternative out-of-sample test developed by Moody's KMV is a *walk-forward* analysis, which is designed along similar lines as the *k*-fold test, except that it controls for the effects of time. The model is estimated up to a certain year and the parameter estimates are then used to score the observations in the *next* year. These model scores are *out-of-time*. The model is re-estimated including one more year of data and we repeat the analysis for the next year and continue until the end of the sample. These out-of-sample out-of-time scores are combined into a single prediction set so that the accuracy ratio and the power curve can be calculated for the combined set. The out-of-sample accuracy ratio is then compared to the corresponding in-sample accuracy ratio and power curve.

No data from a future period is used in fitting the model and only data from future periods is used for testing it. The parameter estimates are checked for stability across the different samples. Figure 9 presents the results from this analysis.

Results

The difference between in- and out-of-sample plots for the walk-forward results is still minor at both the 1- and 5-year horizons in Figure 9. The difference in AR between the in-sample and out-of-sample results is 3% and 1.1% respectively. Furthermore, RiskCalc v3.1 South Africa outperforms Z-score in an out-of-time context at both the 1- and 5-year horizons.[11]

FIGURE 9 Out-of-sample Performance (1- and 5-year) RiskCalc v3.1 South Africa Walk-forward



4.6 Model Calibration and Implied Ratings

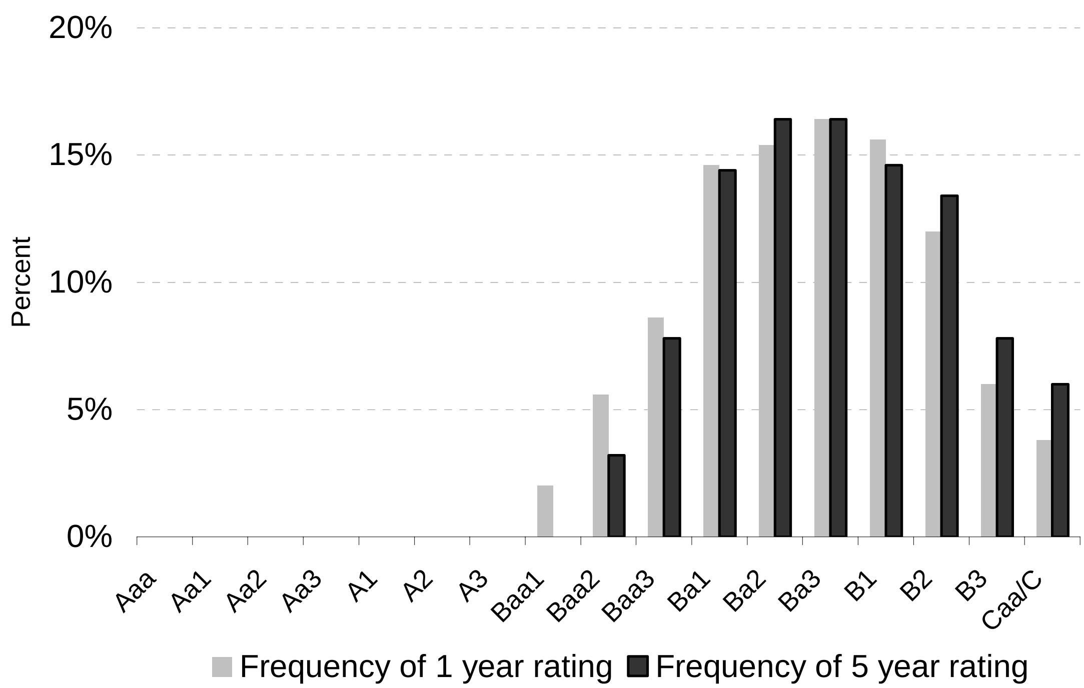
To aid in the interpretation of an EDF credit measure, an EDF value is mapped to a .edf rating (an EDF-implied rating). All RiskCalc v3.1 models to date have used the same mapping. This mapping is designed so that:

- There is a large range of .edf ratings (as required for economic and regulatory applications);
- No one rating contains too many credits (as required for economic and regulatory applications);
- The distribution of the 5-year ratings is approximately the same as the distribution of 1-year ratings (for consistency with rating-based analysis applications);
- The EDF value associated with a .edf rating is approximately the same as the observed historical default rate associated with a Moody's Bond Rating (for consistency with rating-based analysis applications).

[11] The out-of-sample ARs are 42.0% and 37.3% for the 1-year and 5-year models, respectively. These out-of-sample ARs are both 16 percentage points higher than Z-score for the 1- and 5-year models, respectively.

Figure 10 shows the distribution of CRD observations by rating category in the development sample (for the Credit Cycle Adjusted EDF credit measures over the full time period). Note that 10 categories between Baa1 and Caa/C are utilized and that less than 20% of the observations are in any one category. The 1-year and 5-year distributions peak at Ba3. While not reported here, other research has shown that the distribution of the CCA EDF implied ratings changes over time with the credit cycle, while the distribution of the FSO EDF implied ratings remains relatively stable over time.

FIGURE 10 EDF-implied Ratings for the 1- and 5-year Models in the Development Sample



5 CONCLUSION

The Moody's KMV RiskCalc v3.1 South Africa model is built using actual financial statement and default information provided by South African financial institutions.

The power of the model compares favorably to other publicly available alternatives that we have tested. We have demonstrated that the increase in power is consistent across industry sectors and size classifications as well as for different time periods. We have also shown that the power advantage is maintained both in-sample and out-of-sample.

The RiskCalc v3.1 South Africa model controls for differences in the default risk across industries in the FSO mode. In addition, in the CCA mode, it adjusts the EDF level to reflect the current stage of the credit cycle in the given industry. If default risk in a given firm's industry is high, the EDF level is adjusted upward. Likewise, when default risk is low, the EDF level is adjusted downward. This additional feature of the model allows users to monitor their portfolios on a monthly basis.

The Moody's KMV RiskCalc v3.1 South Africa model will be very useful for financial institutions seeking to implement quantitative tools for originating loans, managing portfolio risk, and meeting regulatory requirements. It also provides the user with an objective external benchmark of the risk associated with a private firm, useful in securitizing middle-market debt. Finally, as an established benchmark, RiskCalc v3.1 enables institutions to communicate with one another about their exposures.

6 REFERENCES

1. Altman, E., J Hartzell, and M. Peck. "Future of Emerging Market Flows." New York: Salomon Brothers, Inc, 1995.
2. Crosbie, Peter J. and Jeff R. Bohn. “Modeling Default Risk.” San Francisco: KMV, 2003.
3. Dwyer, Douglas, Ahmet Kocagil, and Roger Stein. “The Moody's KMV RiskCalc v3.1 Model: Next-Generation Technology for Predicting Private Firm Credit Risk.” Moody's KMV, 2004.
4. Falkenstein, Eric. “RiskCalc Private Model: Moody's Default Model for Private Firms.” New York: Moody's Investors Service, 2000.
5. Kocagil, Ahmet E., Alexander Reyngold, Roger M. Stein, and Eduardo Ibarra. “Moody's RiskCalc Model for Privately-Held U.S. Banks.” Moody's Investors Service, 2002.
6. Murphy, Adrian; Ahmet Kocagil; Phil Escott, and Frank Glormann. “Moody's RiskCalc for Private Companies: Portugal.” Moody's Investors Service, 2002.
7. Woolridge, J.M. *Introductory Econometrics: A Modern Approach,* South Western, 2000.
8. Basel Committee on Banking Supervision. "International Convergence of Capital Measurement and Capital Standards (A Revised Framework)." Bank for International Settlements, 2004.