Moody's Investors Service
Global Credit Research

December 2000 **Rating Methodology**

Contact	Phone
New York	
Eric Falkenstein	1.212.553.1653
Andrew Boral	
Ahmet E. Kocagil	

RiskCalc™ for Private Companies II: More Results and the Australian Model

Rating Methodology

Overview

In a continuing effort to provide benchmarks for middle market companies, Moody's has created a model for estimating firm default probabilities using only financial statements.[1] Specifically estimated and tested against middle market firms, Moody's RiskCalc for Canada and the US was released in June 2000. The model allows one to quickly and efficiently attach default probabilities and rank firms from least to greatest probability of default . An unprecedently large dataset of private companies gives us the ability to estimate and validate such a model. As a purely objective, quantitative model, it can serve as a transparent benchmark of credit quality, one that summarizes financial statements into a single number.

This report documents the following:

- Description of Moody's unique private firm database in Australia, with comparisons to the data on US and Canada,
- An updated description of the Moody's methodology for predicting default,
- A comparison of the relationship of various financial ratios to default, and
- Empirical tests of Moody's model.

The following is meant to be a self-contained description of the derivation and testing of Moody's default model, however, some nuances may be omitted. A more complete documentation of the approach is contained in RiskCalc™: Moody's Default Model for Private Companies.[2]

1 "Middle market" means "unlisted" or "private" firms without traded equity information.
2 For further information please refer to http://www.moodysqra.com/privfirm/56402.asp.

continued on page 3

Authors

Eric Falkenstein
Andrew Boral
Ahmet E. Kocagil

Senior Production Associate

John Tzanos

Table of Contents

Introduction

Default modeling is important for the simple reason that defaults represent a statistical expense to lending. These costs can be direct, such as in the expected charge-off rate, or indirect, such as the required capital for lending to firms with various levels of expected charge-offs. While we refer the interested reader elsewhere for the uses of default modeling, a selected list of applications includes the following:

- **Regulators:** Nationally and internationally, bank regulators are looking for a validated metrics in order to assign required capital for internal bank grades.
- **Collateralized Debt Obligations[3] (CDOs):** Moody's is currently a major user of RiskCalc, public and private, in the evaluation of non-Agency rated credits within CDOs.
- **Monitoring:** The efficiency of reviewing hundreds or thousands of obligors is greatly enhanced by focusing upon those 25% that have the 'worst' financial statements, or those that show the greatest deterioration in financial statements. A quantitative tool is useful for this filtering.
- **Decisioning/Pricing:** A good credit model allows one to expedite underwriting for top-tier obligors, price more granularly for mid level obligors, and signal which market segments are not adequately compensating for risk. There are no bad loans, only bad prices.

For each objective one needs a number that allows unambiguous comparison. Further, for pricing, CDOs, and regulatory concerns one needs not only a powerful and efficient tool, but one that is calibrated into a default probability. RiskCalc is aimed at providing a benchmark for all these purposes and more. To qualify as a benchmark, the metric must satisfy four conditions:

1. *It must be understandable*,

   In customer surveys transparency trumps accuracy as a useful model attribute, as users need to understand why a model works, especially a model without an extensive track record. It should be clear what ratios are driving or mitigating a particular final assessment.

2. *Powerful*,

   A model that is not powerful is clearly a nonstarter. One good symptom of the power of risk metrics is that they not only be used for portfolio reporting, but by personnel involved in decisioning and pricing.

3. *Calibrated to default probabilities (DPs)*,[4] and

   without calibrated default probabilities within the output, the model's usage within pricing, capital estimation, and CDOs is unclear

4. *Empirically validated*

   Without documented performance on large out-of-sample datasets, prudence dictates that a model must viewed skeptically. Such testing also gives the user comfort the model was not 'overfit'.

If a model does not satisfy all of these conditions, it may be a useful tool, but it simply can't be a benchmark. A model might be better on three of the four dimensions, but that would still imply it cannot be a benchmark. While we are confident that our model is extremely powerful, we acknowledge that more powerful models exist, the most obvious being models that incorporate subjective, nonstandard, or local information in efficient ways. Nonetheless, RiskCalc is a true benchmark: easy to implement and understand, powerful, calibrated and validated.

The application of RiskCalc to Australia is a natural extension of our base Canadian/US model. Importantly, we see many more similarities than differences among countries in terms of how default probabilities are related to financial ratios. This is useful because it implies that lessons learned in one country can generally be applied to other countries: we do not have to relearn credit analysis for each country. The data presented herein, in addition to unpublished experience with other countries, suggests there are universal credit factors in financial statements, specifically:

1. profitability
2. leverage (gearing)[5]
3. liquidity
4. size
5. inventories and
6. growth rates

---

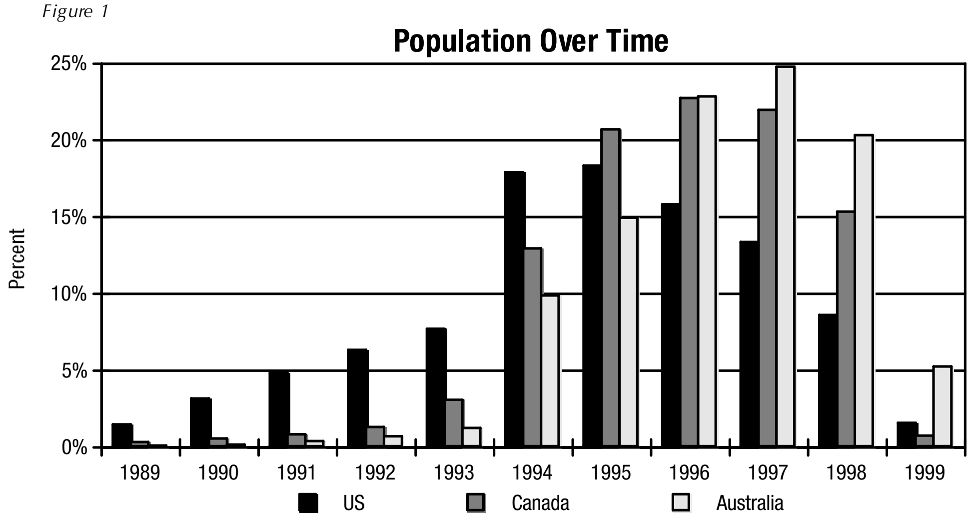
3 CDOs include CBOs and CLOs.
4 We will use the notation DP to represent Default Probability.
5 'Leverage' is an American version of the British term 'gearing', both of which are straightforward metaphors for what is going on. With apologies to Anglophiles everywhere, 'leverage' is our preferred label for ratios related to debt-to-net worth.

These risk factors have a similar relationship with default probabilities around the world, and graphs of the relationship only differ in the steepness of their slope and curvature—there are no countries, for example, where higher leverage is associated with lower default probabilities. While refinements in the various countries can differ, these factors are the backbone to any summarization of a financial statement. As will be shown below, the largest difference between the Australian and Canadian/US models is not so much in the relative weightings on the financial ratios, but instead in the mapping of the output into default probabilities. Statistical power is conserved across countries far more than is the absolute level of DP (default probability) generated.

Data Description

Figure 1 shows that for Australia, Canada and the US most of our data is from 1994 onward. This implies that we miss the most recent recession in all these countries.[6] This is the result of the newness of our initiative and the relative ease of retrieving more recent records. We are constrained by the data we currently have, and we simply can't go back in time and get more data from the last recession, yet clearly default prediction that only works during good times is of limited interest.

Figure 1



Most statements are from the past 6 years

We are confident in our model's relevance in recessions for two reasons. First, our default probability mappings are from population default probability estimates. These are estimates of default probabilities through the cycle, which means abstracting from whether we think we are in a recession or not. Thus, the abnormally low default probabilities of a post-recession period are not assumed to be representative of the future. Secondly, we are confident the relative weightings of risk factors like profitability and leverage are robustly estimated even within nonrecessionary years. We use US public company data to test this hypothesis, and indeed find that the relative weightings of a probit model estimated in nonrecessionary years is approximately the same as for a model estimated upon recessionary years. Test results for RiskCalc in recessionary times are in the Testing section (page 22).

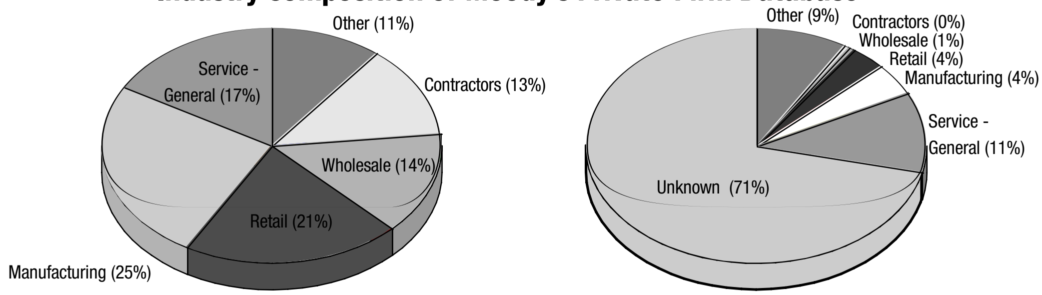
Figure 2 shows the relative industry concentrations. Note that in our Australian database by far the largest category is 'unknown', or 'unreported'. For those companies that do have industry information, the most common designations are services, retail trade and manufacturing. This omission of industry classifiers reflects an important fact of middle market lending. Not only is it sometimes difficult to assign a company to a precise industry code, occasionally, *mea culpa*,[7] the software used to warehouse financial data does not facilitate gathering and storing this information. For whatever reason, the company's financial information is often spread within bank electronic databases as 'general' or 'unknown'. A model needs to be robust in the face of this reality of middle market lending because a model that relies crucially on industry benchmarking will have modest coverage.

When building quantitative models, it is important to recognize as a restraint the imperfection of middle market data. A model must work with the data available, the quality of which invariably shocks people who are new to this area. A model made using an imaginary, perfect set of information would be quite different from what is presented here, though since we have never seen a perfect set of data such benchmarking is pointless.

6 The last official recessionary periods for Australia, Canada and the US were in 1991.
7 Moody's is an international leader in spreading, or deal capture, software solutions.

Figure 2

Industry Composition of Moody's Private Firm Database



Australian statements have significantly fewer industry classifications, though of those noted, most firms were in services, trade, and manufacturing

The ultimate payoff is the number of unique defaulting, and, to a lesser extent, non-defaulting companies. Given the fact that most companies do not default, defaulting companies are more rare and thus more valuable from an information perspective. Much of the dearth in default data is due to the vagaries of data storage within financial institutions. Defaulting companies are often purged from the system after their troubles begin, which creates a sample bias in that the default probability implicit in current bank databases is invariably low, even for a non-recessionary period. Default is defined as any of the following events:

- 90 days past due
- bankruptcy
- placement on internal non-accrual list
- write-down

If we determine any of these actions being taken we consider the obligor defaulted as of that date. If a company has its loans restructured in such a way that there is an adverse effect upon the lender, such as moving payments back in time without any compensation, in general we do not capture this as a default. This is not our intention, as adverse restructurings are part of Moody's corporate definition of default. It is a limitation of the data: rarely are such restructurings properly recorded in internal systems. If we can capture such 'adverse exchanges' they would be classified as defaults, yet in practice it is more accurate to simply say that we do not, generally, capture such events as defaults.

Table 1 illustrates the different samples used in our private firm RiskCalc estimation. Note that there are approximately as many firms and statements in Australia as in the US. Gathering data in the US is made more difficult by a more fragmented banking community, and the lack of a centralized repository as many European countries have. In the intermediate future we expect the US to lag other nations in the amount of private firm data its banks can provide.

Table 1

Information on Sample Data

Country	Time Span	Unique Firms	Unique Firm Defaults	Financial Statements
United States Private	1989-99	33,964	1,393	139,060
Australia Private	1990-99	27,712	1,447	79,877
Canada Private	1991-99	4,472	271	18,538

In addition to the number of 'goods and bads' we also report the number and quality of statements in **Table 2** below. One can see that approximately one-third of our US and Australian data consists of incomplete statements, primarily due to missing statements in prior periods. Thus, we mainly are missing the growth information, which necessarily requires the prior years' financial statements. Only a minority of the statements in any country is audited, which adds uncertainty in that unaudited statements are less accurate.

Table 2

Breakdown of Financial Statements

Country	Total Financial Statements	Incomplete Statements	Audited Statements
United States Private	139,060	41,383	47,823
Australia Private	79,877	39,966	23,697
Canada Private	18,538	4,790	4,524

Aggregate Default Profitability Assumptions

Before going into the actual ratios, it is useful to first describe some higher level differences among the countries. First, taxes. Though exceptions are numerous, in general there is an incentive to pay out profits to a proprietor as salary. Dividends are taxed twice (at the corporate level and again as dividends), salary only once (as income), so there seems to be an incentive in each country to pay out as much as possible in salary to avoid a second tax.[8] Anecdotally, middle market lenders in all three countries have identified this 'gaming' of the system as potentially relevant, and no more or less so in any of them.

Table 3 illustrates two indicators of bank credit quality that help guide our top-down opinion of aggregate default probabilities within these countries. First, in general the banking sectors in the various countries have roughly similar ratios of nonperforming loans as a percent of total assets (column 3), which implies that asset quality is approximately equal across the countries. The different proportions of consumer and small business lending in these figures clearly make aggregate comparisons difficult, nonetheless they are still informative. Another top-down note within Table 3 is the median P/E of banks, especially the P/E ratio of banks relative to other corporates in each country. The ratio is highest in Australia, then Canada, and lastly the US. This fact taken alone would seem to imply that the riskiness of Australian banks is lower than of Canada or US banks. Taking this further we note that the average P/E[9] for all corporates in Australia is generally lower than for Canada and the US, so that, relatively, banks in Australia are a less risky investment than in Canada or the US.

Lastly, we performed an informal survey of data contributors, and invariably the expected charge-off rate, the same one used in provisioning and pricing models, was somewhere between 30 and 50 basis points for middle market lenders in each country. Loan recovery rates, also from informal survey, as well as what is documented in Gupton (2000) and Falkenstein (2000), are approximately 70-80%. When combined with a US loan recovery point estimate of 75%, this again gives us comfort that the Canadian and Australian experience is quite similar to the US experience. These survey points suggest a reasonable default probability boundary between 1.2% and 2.0%.[10] In sum, it appears most appropriate to assume that both Canada and Australia have a 1.7% default probability, the same rate we assume for the US. (As with all references to the US market, refer to the RiskCalc Technical Document for further information.)

Table 3

Indicators of Bank Credit Quality

	Average Local Index P/E	Average Bank Local P/E	NPA/Total Assets 1995-99 Avg.
Australia	20.8	14.9	0.70%
Canada	33.3	12.3	0.68%
Large US Banks	26.9	11.5	0.88%

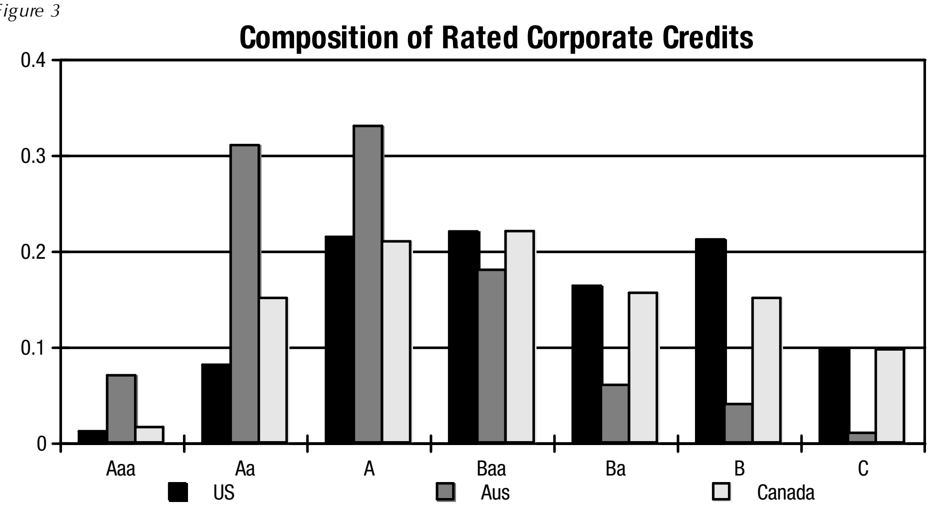
Indices underlying the P/Es were the All Ordinaries, TSE100, and the S&P500, for Australia, Canada, and the US, respectively. P/E ratios were computed in early October, 2000.

Top-Down data on asset quality and stock market valuations suggest Australian bank risk is no greater than in the US

Finally, it should be noted that the average Agency rated corporate credit quality within a country has limited significance for generalizing about their unrated credits. Outside the US there is a selection bias for Moody's ratings, in that firms with higher credit quality tend to find the value of a Moody's rating most valuable, and thus companies within Australia tend to have higher ratings than do companies in the US. Over time we expect Moody's credit composition in Australia to migrate towards the US distribution, in the same way that Canada's corporate rating distribution has begun to more closely approximate the US. This is all due to reasons outside the general credit quality in each country, and instead reflects the legacy of Moody's US origins.

8 The individual tax rates tend to be higher than corporate tax rates in all countries, though not enough so to offset the effect of the double taxation of dividends.
9 Price/Earnings ratio.
10 Using the fact that charge-offs=default rate x (1 - recovery rate).

Figure 3



While Australian companies tend to have higher ratings than the US, this is most likely due to a selection bias

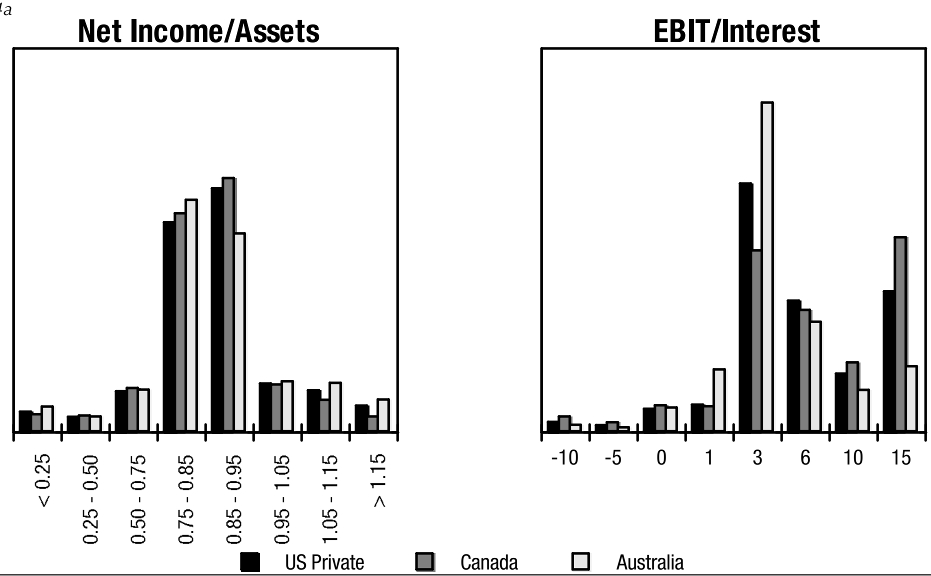
Comparitive Analysis: Distributions

The most obvious first-order adjustment to make when moving outside the US is to correct for accounting variations: depreciation allowances may be quicker in certain countries, relative tax rates may encourage more dividends, etc. Much of the effect of these differences can be addressed simply by normalizing by the native distribution. For example, the median leverage ratio measured as liabilities/assets is 62% in the US and Canada, while it is 80% for Australia. It would seem appropriate to consider a leverage ratio 'high' because of its standing within its country, as opposed to its standing when compared to ratios of companies in other countries, as differences will be due not only to accounting, but to business volatility that varies from country to country. This sort of adjustment is implicit in peer comparisons, whereby credit analysts have long adjusted accounting ratios within different industries on the assumption that what's low in retail may not be low in manufacturing.

Below we document the different distributions of ratios used within RiskCalc for each of the countries. For the US this data only pertains to private companies.[11] It allows us to highlight several key issues in the data. When combined with the information on default probabilities (using the univariate default relation, page 19, below), it helps us zero in on distinctions and similarities between the country's financial statements.

Profitability: Australia, Canada and the US all have roughly similar profitability, as measured by Net Income/Assets, or EBIT/interest.[12] Australia's EBIT/interest is more centered, while Canada's less so, but these differences do not appear to have great economic significance.

Figure 4a



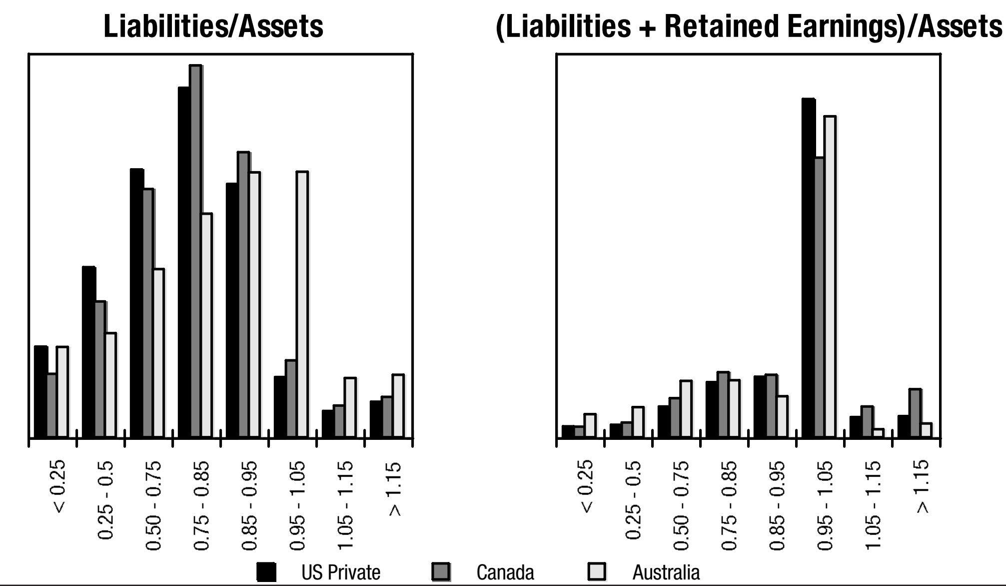
11 Specifically, not Compustat data.
12 EBIT is 'Earnings Before Interest and Taxes'.

Medians

	Australia	Canada	US
NI/A	4.0%	4.2%	4.5%
EBIT/Interest	1.76	1.73	1.82

Leverage/Gearing: At first glance, Australia appears to be significantly different in its leverage/gearing from either the US and Canada, where the latter two appear quite similar. Australia appears to have the significantly higher leverage. For example, 20% of all Australian companies have negative net worth (L/A>1), while only 6% and 8% have negative net worth in Canada and the US, respectively. Looking more closely, however, we see that this is primarily due to retained earnings, which are close to zero in Australia, but around 25% of assets in Canada and the US. If we abstract from Retained Earnings, the distribution of leverage is quite comparable across the countries. Given that we estimate the expected default probabilities to be comparable across all three countries, it appears that the lack of retained earnings, the lack of this "buffer" within the balance sheet, does not affect the financial strength of firms in aggregate, though it does affect the financial strength of firms cross-sectionally. This result remains puzzling, and is an outstanding research item.

Figure 4b



Medians

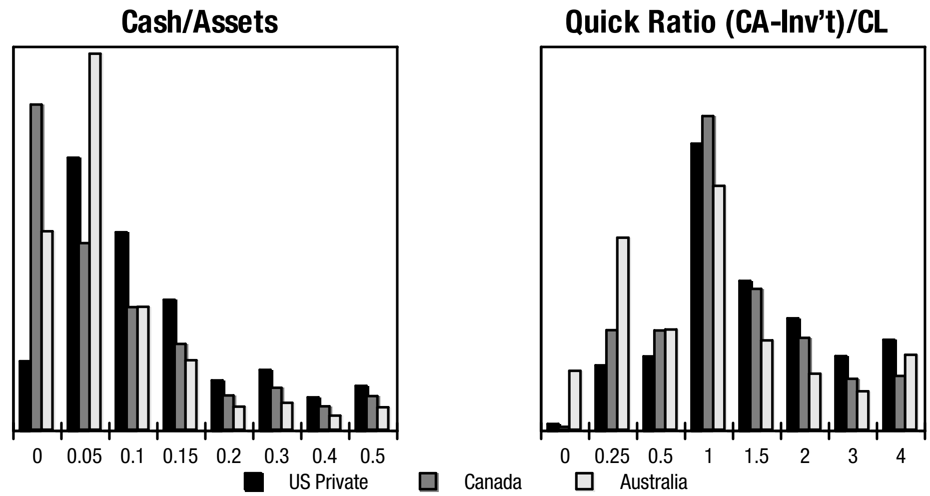
	Australia	Canada	US
Liab/Assets	83%	68%	63%
Ret.Earn./Asset	0.3%	23%	28%
(Liabilities+Ret.Earn.)/Assets	99%	98%	98%

Liquidity: Looking first at cash/assets, Canada sticks out, with 40% of all Canadian companies having zero cash/assets. This is primarily the peculiarity of a competitive equilibrium where many sweep accounts are used.[13] While this practice exists in all countries, and is not legally encouraged or discouraged in any country, Canadian banks make the arrangement more often.

In the quick ratio, however, Australia comes out with the weaker balance sheet. The quick ratio measures current assets minus inventories divided by current liabilities, and thus is a measure of the 'quick' assets available relative to the liabilities due soon. In addition, excepting the large number of 'zeros' for Canada (which admittedly is a big exception), Australia has lower cash/asset levels.

13 In a sweep account deposits are used nightly to pay down loan balances, and in the process ensure that 'cash and equivalents' are zero. Cash in this context includes marketable securities.

Figure 4c

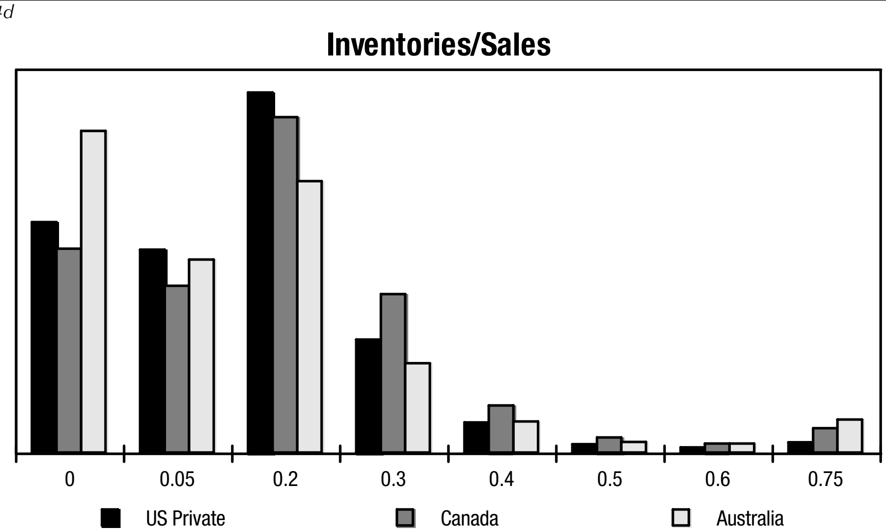


Medians

	Australia	Canada	US
Cash/Assets	0.5%	0.9%	4.0%
Quick Ratio	56%	78%	95%

Inventories: Australia has many more 'zero' observations for inventories, which would, on first glance, imply a lower risk, yet unlike in Canada or the US, firms with zero inventories in Australia have higher default probabilities than those with only a marginal amount. In Canada and the US the zero inventory figure acts as a dummy variable for service industry firms, who tend to have lower risk. It is our best estimate that the reason why Australian firms have zero inventories so frequently is simply because inventories are often mistakenly recorded as zero when they are really 'unknown' — thus firms with zero inventories tend to behave more like the average firm as opposed to service companies.

Figure 4d

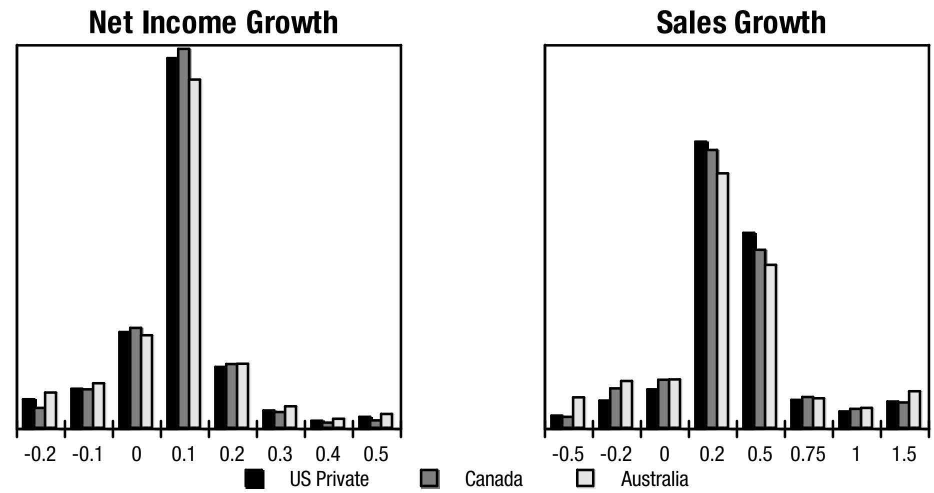


Medians

	Australia	Canada	US
Inv't/Sales	1.5%	7.5%	4.2%

Growth: Growth measures are very similarly distributed among the countries.

Figure 4e

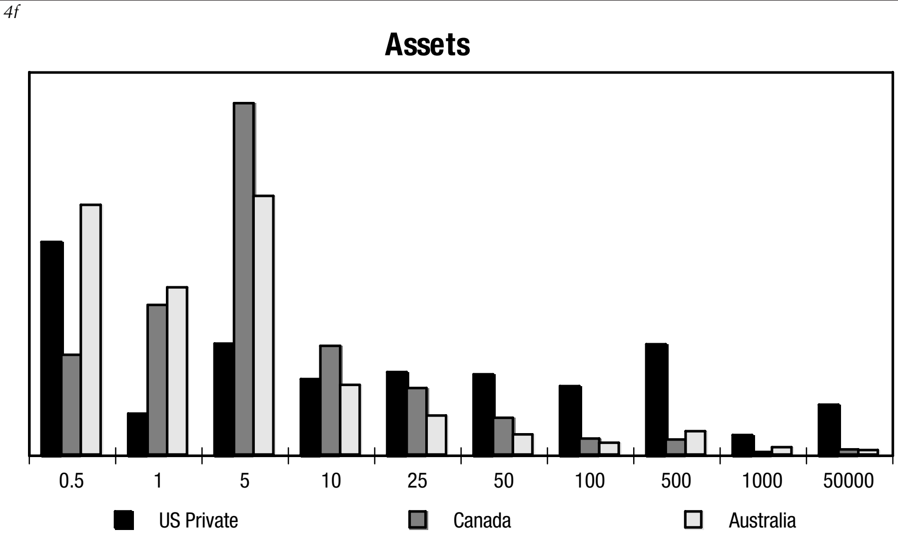


Medians

	Australia	Canada	US
Sales Growth	5.6%	6.2%	7.6%
NI Growth	0.0%	0.4%	0.5%

Size: Size is a notorious correlate with various inputs, most significantly the quality of financial statements and our measurement of default. Larger companies tend to have audited statements that are of better quality. More importantly, perhaps, is that our measure of default is more accurate for the larger firms. While we do our best to make sure that companies in our database are truly defaulted or nondefaulted companies, and in fact exclude more data than we use because of this effort, inevitably we do make some misidentifications. Therefore, size and data quality correlate positively. Since models work better on firms with cleaner data (see RiskCalc technical document), a corollary of this is that models work best on our Canadian dataset and worst on our Australian dataset, primarily because of the average firm size for each sample.

Figure 4f



Medians

	Australia	Canada	US
Assets ($Millions)	0.740	1.496	2.021

Model Description

Modeling default quantitatively is often counterintuitive, in that financial ratios, no matter how bad or good, never suggest a company will probably fail.[14] To give an example, of all the firms both losing money and insolvent, their default probability is still relatively infrequent, though it is about three times higher than for firms that don't meet these criteria. Financial performance as reflected in accounting ratios affects default probability the way that the speed of a car affects the probability of an accident: there's a strong , highly nonlinear relationship, yet crashes remain infrequent at any speed. This is reflected in the model's inner workings, in that even for extremely poor ratios, the model will still generate relatively low default probabilities (<20%).

The first step in any model is the selection of input variables. With at least a hundred potential financial ratios in standard accounting textbooks, the combinations are astronomical. Clearly there is no feasible mechanism for looking at each combination of ratios, and the result of an exhaustive step-forward approach would also produce many highly significant combinations purely by chance. Thus we constrained our approach to use only common-sensical relations, and targeted the big 6 categories of profitability, leverage, liquidity, size, inventories and growth. Within these categories we made sure that ratios with 'wrong' signs were excluded. The result, was a set of 8 ratios, and size, that we used in RiskCalc.

The modeling approach can be briefly summarized in three steps: transformation, modeling, and mapping. It is a nonstructural model, in that it does not use an explicit function based on theory, but it is highly informed by the collective experience of Moody's. There is a trade-off between in-sample fit and out-of-sample robustness, and our bias is towards the simplest functional form and the smallest number of inputs. All of our inputs fall into well-known correlates with financial distress, such as profitability and leverage. The three steps to our modeling procedure are described below.

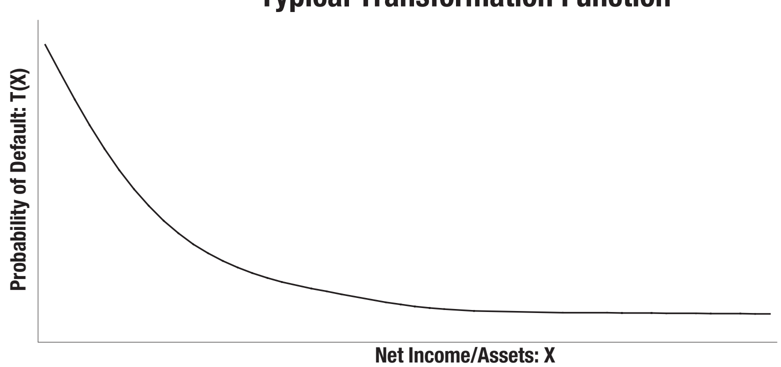
Transformation

The first step involves mini-modeling the input ratios by replacing the input ratios with an estimation of its corresponding univariate default probability over 5 years. That is, each ratio, such as net income/assets, is related to a default probability; instead of using the input ratio, we use the default probability corresponding to this ratio. This captures much of the nonlinearilty of the problem, normalizes the inputs to a common scale, and allows one to monitor the marginal effect within the model simply by observing the univariate default prediction.

The chart below shows a transformation function used within the model, in this case it approximately reflects the relation Net Income/Assets to default rates. The transformation turns a ratio like Net Income/Assets that ranges from -0.25 to 0.50 into a number between 0.02 and 0.08, or a number like sales growth, which ranges from -0.75 to 2.5 into a number between 0.02 and 0.06. This latter range reflects the five-year default probability for firms based on only using this single ratio. In the example below, the transformation implies a significant risk difference between the least profitable and the moderately profitable companies, while in turn there is little difference between companies with very high and moderate profitability; this is what is meant by a 'nonlinearity'.

Figure 5

Typical Transformation Function



Transformations both illustrate how ratios affect the model and capture nonlinear relationships

14 By "probably" we refer to a probability that is in excess of 50%.

Model

The second step involves using the transformed inputs within a multivariate model so that the weights assigned to the multivariate model are appropriately adjusted not only for their univariate power, but for their power in the presence of other, often correlated information. Thus the model accounts for correlations, just not through any direct interaction terms such as net income x sales growth.[15]

We then use these transformations as the input to a binary model that predicts default. In our case we estimate a probit model, which uses the normal or Gaussian cumulative distribution function, specifically:

$$y = Prob(default \mid x; \beta, \sigma) = F(-\beta' T(x)) = \int_{-\infty}^{\beta T(x)/\sigma} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{t^2}{2}\right) dt \qquad (0.1)$$

The advantage of the probit model, as opposed to, say, ordinary least squares, is that it specifically accounts for the fact that the output, being binary, is restricted to between 0 and 1. It is a more statically efficient method for estimating these problems because when the error is asymmetrically distributed simple ordinary least squares is inefficient.[16]

The resulting model once estimated is a generalized linear model, in that it is a nonlinear function[17] of a linear model: $y = \Phi(f(x,\beta))$, where the linear part is simply

$$f(x,\beta) = \beta_0 + T_1(x_1)\beta_1 + T_2(x_2)\beta_2 + \ldots + T_{10}(x_{10})\beta_{10} \qquad (0.2)$$

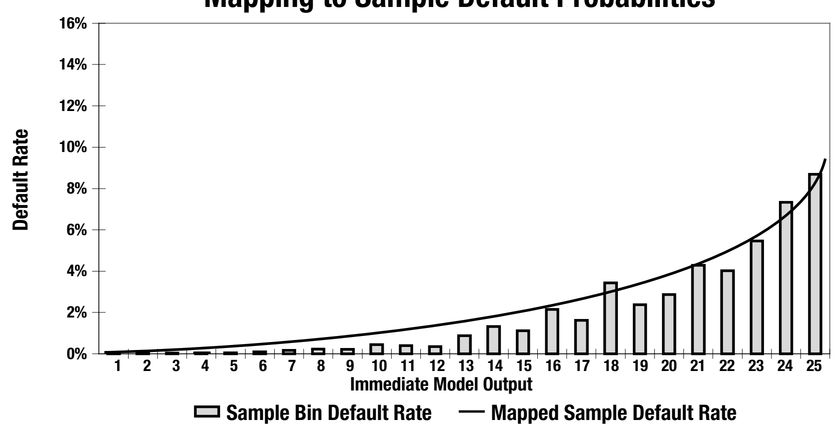
the T(X)'s are the transformations (as in **Figure 5**, above).

Map

The final part of the modeling, mapping is similar to the transformation of ratios. We take the output of the probit model and find the best fitting function that maps the output into the sample default probability. This is done because invariably the output from the probit model tends to overestimate the true probability within sample. It is a common problem in applied probit or logit prediction, and relatively straightforward to correct. **Figure 6** below shows how one takes the output of the model and maps it to sample default probabilities. Note that the ordinal ranking along the x-axis implies that the output of the probit model could be in any units, it simply does not matter. One estimates the relation between model output and sample default probability using a variety of smoothing algorithms, which on a univariate problem are not meaningfully different. This smoothing process is done identically to how we smooth our input transformations.[18]

Figure 6

Mapping to Sample Default Probabilities



Model output is calibrated to Sample Default probabilities via a smoothing algorithm

---

15 We are not categorically against such extensions, but we are very wary of the degrees of freedom they bring forth. Moreover, there are no obvious interactions within the data, and the number and type of nonobvious interactions is sufficiently numerous to introduce more error from overfitting than nuanced enhancement at this point.
16 For example, for the highest annual default rate imaginable, 30%, this implies a 0.7 error, nondefault a 0.3 error.
17 Specifically, a sigmoidal function.
18 We now use the Hodrick-Prescott filter, though virtually any smoother that retains the exponential nature of the extremums will do.

Finally, we have to adjust the *sample* default probability to our projected *population* default probability. The sample yields a biased default probability estimate for two reasons. First, it was estimated during primarily nonrecessionary years for all countries, so the default probability in nonrecessionary years is lower than average, over good times and bad. Second, default identification is sufficiently difficult that we invariably undercount defaults. The net result is that our sample 1- and 5-year default probabilities are nowhere near our best estimate as to the average default probability. We force our sample rates to equal our population rates by simply multiplying by a constant so that, on average, over the entire sample we had at construction, our average firm had 1 and 5 year default probabilities of 1.7% and 6.8%, respectively.

The ultimate adjustment is simply

$$DP_{pop,bucket} = DP_{sample,bucket} * \frac{\text{population default rate assumption}}{\text{sample default rate}} \qquad (0.3)$$

In sum, the transformation normalizes the input ratios and captures their nonlinearity in a transparent way. Applying a standard binary model to these transformed inputs creates outputs that must be mapped to default probabilities the same way that mortality tables are assigned to people of a certain age: based on historical data. The final touch is the adjustment for the top-down default rate for the entire country, which is usually arrived at outside the model: transform, model, map.

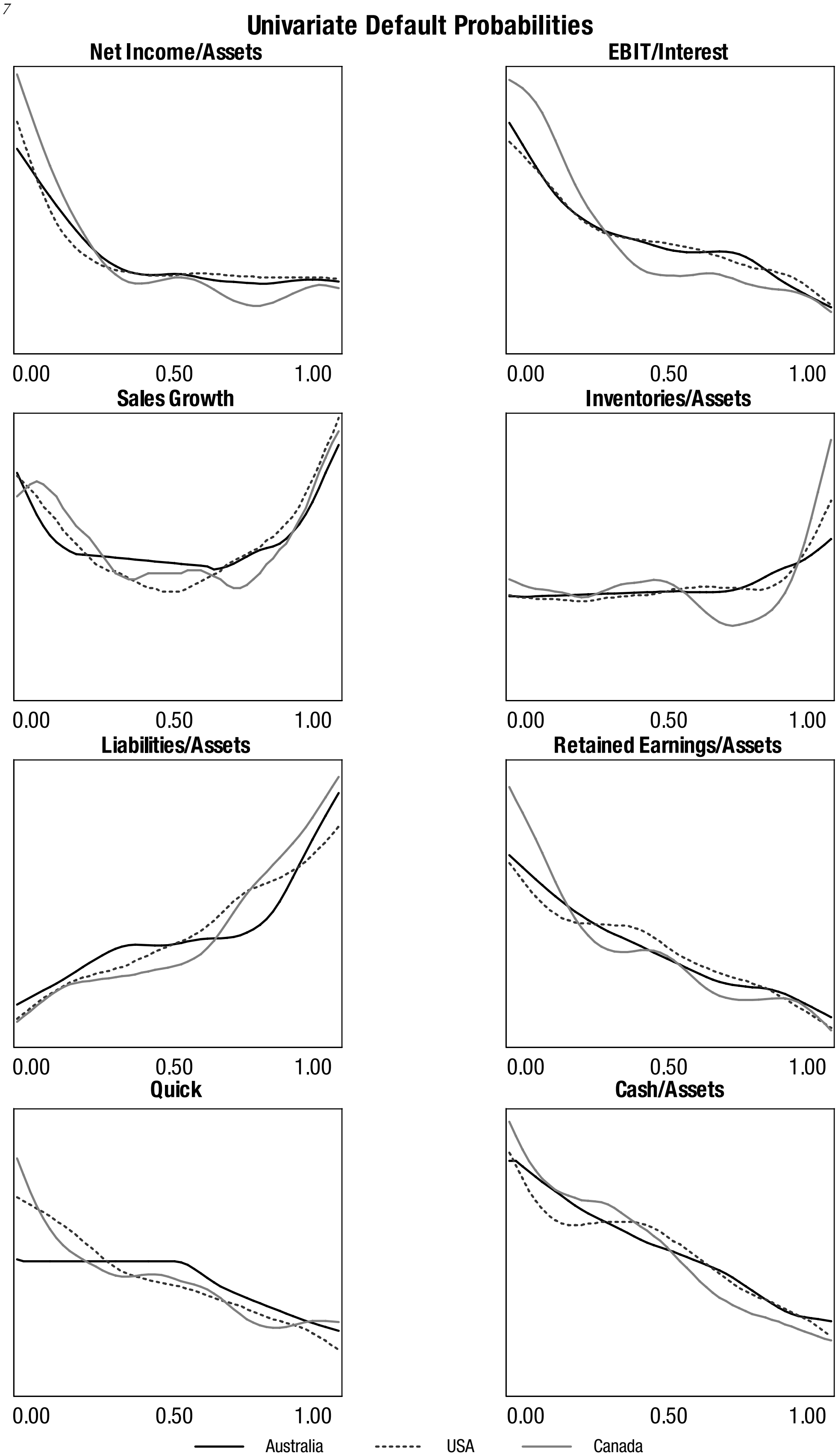
Accounting Ratios And Their Relation To Default

For the US and Canada the relation between accounting ratios such as Net Income/Assets and default probabilities is quite similar. This is also reflected in the model itself, where the inputs are generally treated the same whether or not they are from Canada or the US, the only difference being the zero point for Cash/Assets. In both countries a zero value for cash is a bad signal, but it is significantly a worse one in the US where it is less frequent. Other than this difference (when cash=0), the Canadian and US models are identical.

Figure 7 illustrates the difference between the Australia, Canada, and the US. The main highlights are as follows:

1. Inventories: Higher inventory levels imply higher default probabilities in all countries. The effect is nonlinear, with higher default probabilities accelerating for firms with the highest inventory levels.
2. Profitability: Higher profits, measured as NI/A or EBIT/interest, lowers default probabilities in all countries. The effect is more nonlinear for NI/A, suggesting that EBIT/interest could be more useful for assessing credit quality in intermediate areas, while NI/A may be a better profitability benchmark for those 30% of firms in the lowest quality spectrum.
3. Leverage: Higher Liabilities/Assets ratio implies higher default probabilities in all countries. This relation is more nonlinear in Australia, reflecting a greater acceleration of default probabilities as firms move into higher percentiles. While the lack of retained earnings/assets appears to not affect Australian default probabilities in aggregate, cross-sectionally it does help predict default probabilities in each country.
4. Liquidity: Lower liquidity implies higher default probabilities in all countries. The effect of liquidity ratios is less linear across percentile space for Australia, as the lowest liquid firms do not differ as dramatically from those with average liquidity. For North America, lower liquidity, using cash or current assets, is an almost linear indicator of default risk for both the US and Canada.
5. Growth: For both sets of counties, both high and low growth are associated with higher default probabilities. Sales growth was much less predictive in Australian than in Canada/US data, though this could be more of an issue of data quality than a real difference. (Fewer contiguous statements exist in our Australian database.) For Australia, higher growth firms did not correlate with higher default probabilities. In contrast to US & Canada on experience.

Figure 7



5-Year Cumulative Default Probabilities As Sorted By Various Univariate Ratios

The univariate relationships suggest that, in general, models fitted to Canada or the US data will be powerful in Australia: firms with higher profitability, lower size, lower leverage, etc., have lower default probabilities in Australia just as they do in the US and Canada. Yet we still can improve model performance by making certain adjustments. The process of adjusting the transformation of the input ratios using the Australian experience simultaneously adjusts for different mean ratio levels (such as the fact that Australian mean leverage is 82% vs. 60% in the US) and for differences in the linearity of various ratios (e.g., the greater nonlinearity of the Australian leverage risk factor).[19]

In a multivariate model, however, the univariate relationships to default are not all that matter. Financial ratios are invariably correlated with each other, and these correlations affect the ultimate weightings of different inputs within a multivariate model.[20] A country with exactly the same univariate relationships to default may have different optimal factor weightings due to different correlation of size and liquidity, or profits and sales growth. Hypotheticals aside, the actual correlation between input ratios was similar in the three countries. The two sets of ratios are illustrated in **Table 5** illustrate this. We used the correlation of the transforms, in part because it is most relevant to how it affects the model (given that the model uses the transforms directly in the algorithm), and also because the transformation process removes the outliers which tend to obscure the true correlations. We see that profitability and leverage are generally positively correlated: firms with 'good' profitability tend to have 'good' leverage. Another set of ratios—sales growth and the quick ratio—shows virtually zero correlation in all of the countries.

Table 4

Correlation of the Transformation of Key Input Factors

	NI/A~L/A	Quick~Sales Growth
United States	0.33	0.01
Canada	0.26	-0.05
Australia	0.29	0.03

The correlation of key input ratios is roughly consistent across various countries

By examining the univariate relations above between input ratios and default probabilities, one can avoid much of the minutia of different tax codes and accounting schedules to create powerful and calibrated models in these countries. That is, we do not have to know why a relationship becomes more nonlinear or less powerful, only that it does, and in the process appropriately adjust for local variation.

Weights

One way to understand how the model works is to consider the approximate weightings on various factors. If we group the inputs into their obvious counterparts we get the following:

Table 5

RiskCalc Relative Weights of Risk Factors by Country

	USA	Australia	Canada
Profitability	23%	15%	24%
Leverage	21%	35%	22%
Liquidity	19%	33%	15%
Size	14%	5%	15%
Growth	12%	5%	12%
Inventories	12%	7%	12%

All models consider the same basic risk factors, though the weightings are different

Note that for Australia, growth and size become relatively less important. This is mainly due to the poorer quality data for the Australian dataset, as it contained relatively more smaller companies that have measurement problems with both inputs (financial ratios, especially prior year's growth rates) and outputs (defaults for smaller firms are not tracked as closely for these companies).

19 The transformation process, or the display of univariate default relationships (which in RiskCalc are the same thing), maps the data first from ratio into percentile and then into default rate for illustrative purposes only. The smoothing function that determines the relationship does not use the intermediate step of percentiles, it is only there to provide the user with additional information, and the reader with a more transparent graph.

20 It is a common misconception that correlation between explanatory variables invalidates their use in a model. While correlation does increase the standard error of the coefficients, most modeling techniques remain unbiased, and with appropriate consideration for overfitting, can be efficient.

It should come as no surprise that financial ratios tend to have similar influences on relative risk in different countries. While there are no laws of finance that define universal constants, such as leverage ratios of 90% being associated with default probabilities of 5%, there do appear to be laws of finance that define universal tendencies, such as that higher profitability implies lower default probabilities. This reinforces our confidence in the ability of financial statements to rank order firms from most to least likely to default.

Empirical Tests

The primary testing tool we use for assessing statistical power — the ability to rank-order defaulters and nondefaulters — are power curves. The curves graphically illustrate the ability to exclude defaulters for arbitrary cut-off points, and can be aggregated into a single statistic that allows for numerical comparisons among models.

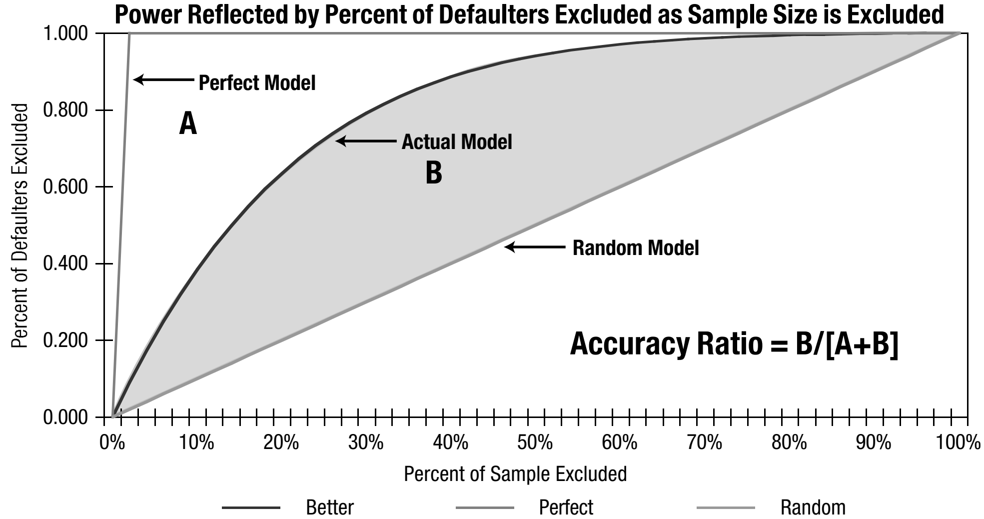
The power curve itself can be defined as follows. It maps the fraction of all companies with the worst score (horizontal axis) onto the fraction of defaulting companies within that group (vertical axis). Ideally, if the sample contained 10% defaulters, then a perfect model would exclude all those defaulters at 10% of the sample excluded: the lowest ranked companies would include all the defaults. Purely uninformative and perfectly informative models are illustrated in **Figure 8**.

In reality defaulters will not be perfectly discriminated, creating a curvilinear function: at 10% of the sample excluded 30% of the defaulters would be excluded, at 20% of the sample 40% of the defaulters would be excluded, etc. This creates a line that is bowed out towards the upper left (Northwest) of the chart: the greater the bow, the better.

Figure 8

An Illustration of the Power Curve

Power Reflected by Percent of Defaulters Excluded as Sample Size is Excluded



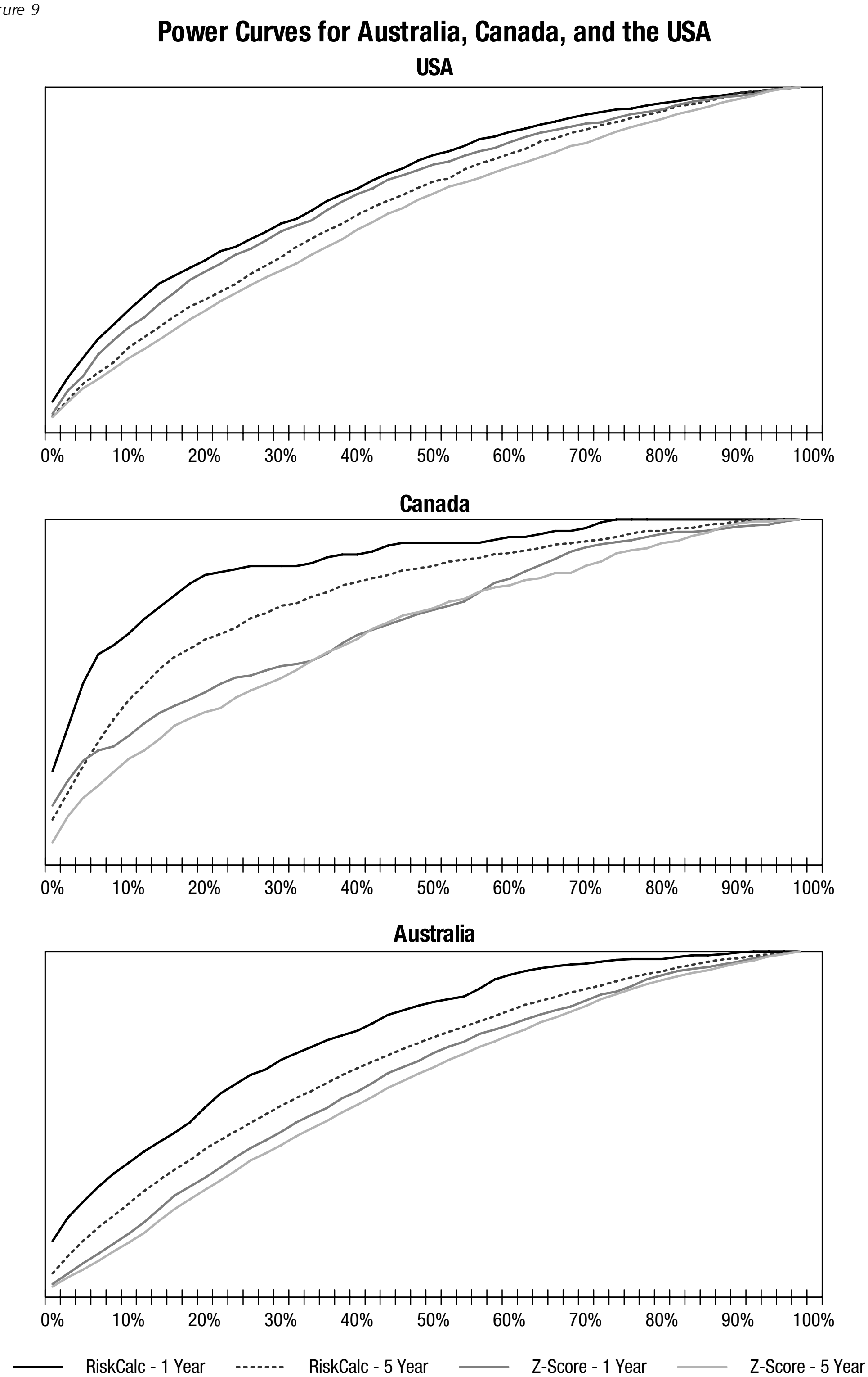
There is also a strict correspondence between the default frequency curves and the power curves (see Appendix), so that the above graphs showing the relationship between default probabilities and a firm score correspond to power curves, and also any statistics derived from power curves. Thus, when looking at a simple graph that shows default probabilities by any metric (as in figures 5, 6, and 7 above), one has to remember that this information implies particular power curves.

It is important that the models predict, not simply explain, and thus we need to test the models using data received prior to the default date. Given the average lifetime of a new loan is 4-5 years, using financial statements several years prior to the default date is extremely useful. This creates a problem because if we simply use each financial statement as an observation, we will double and perhaps triple count defaulting firms: once for each statement prior to the default date. Our method of testing accommodates this complication by going backward in time from the default date, as opposed to forward in time from each statement date. This ensures that each failed firm is counted only once and does not bias our results. The details of the procedure by which we estimate both 1 year and 5 year default probability curves is described in the appendix.[21]

21 Actually they aren't strictly 1 and 5 year default probabilities, they are 'short run' and 'long run' default probabilities, which are approximately of 1 and 5 year duration. Please refer to the Appendix for further details.

For the Canada, Australia and the US, the power is reflected below.

Figure 9



The power can be displayed numerically as accuracy ratios, which represent the area under the power curve as a percent of the area under the perfect power curve. The range is from 0 to 1, 0 being totally uninformative, 1 being perfectly informative.[22]

22 *See definition in figure 8, or refer to Sobehart, Keenan & Stein (2000).*

Table 6

Accuracy Ratios by Model, Country, and Horizon

(Standard errors in parentheses)

Country	Model	1-2 Years	5-3 Years
US	RiscCalc-US	0.537 (0.018)	0.333(0.018)
US	Z-Score	0.420 (0.018)	0.240(0.018)
AUS	RiscCalc-US	0.368(0.014)	0.279(0.014)
AUS	RiscCalc-AUS	0.397(0.014)	0.308(0.014)
AUS	Z-Score	0.277(0.016)	0.199(0.016)
AUS	RiskCalc Cross-Validation	0.374	0.322
CAN	RiscCalc-US	0.577(0.036)	0.368(0.038)
CAN	RiscCalc-CAN	0.583(0.036)	0.380(0.038)
CAN	Z-Score	0.497(0.036)	0.341(0.038)

The higher the accuracy ratio the better. All models display significant power, and RiskCalc is a significant improvement on the Z-score benchmark

Model performance is a function of two factors: the data and the model. Thus the performance of a model should only be compared with the performance of another model on the identical dataset, as only this allows for a truly apples-to-apples comparison. Thus, above, the groupings were made to allow for the direct comparison of the various models within each particular country. The power at predicting longer horizons is less than the power at predicting shorter horizons. It is easier for any model to see 1 year ahead rather than 5 years ahead. This does not mean that the 1 year forecast is more important, however, and thus we produce both 1 and 5 year forecasts

Table 6 makes several major points. In all cases RiskCalc dominates Z-Score, a benchmark chosen for its popularity in accounting and CFA texts.[23] While other models make more relevant comparison, we are constrained to using publicly available models.

Applying the US model to Australia provides a truly out-of-sample test of the US model, in that the Canadian/US version was estimated prior to examining the Australian data. We see that the US model as applied to Australia is significantly less powerful than the native Australian model, as would be expected given that the Australian model was fit to Australian data. But the difference is not severe (0.308 vs. 0.279), especially in relation to the Z-score benchmark (0.192). Similarly, if we ignored the RiskCalc distinction for Canada, we also see very little diminution in power.

Another out-of-sample test was the 'RiskCalc Cross-Validation' applied to Australia. This is another way of estimating the standard error of the RiskCalc approach that is more robust than simply approximating the standard error for a single model on a given set of data. The cross-validation approach uses 80% of the sample to estimate the model, and 20% to test. By sequentially choosing different 20% hold-out samples, we can generate a cumulative set of forecasts for the entire dataset, and then examine the performance of this approach. It is not a test of RiskCalc-Australia directly, but instead the general methodology, and gives a good indication of its robustness. We see that for the short-run sample power degrades somewhat from the in-sample power (0.374 vs. 0.397), whereas for the longer horizon power actually improves (0.322 vs. 0.308). The decline in power for the out-of-sample model is modest, understandable, and still suggests great improvement over the Z-score benchmark. The increase in power at the longer horizon is, at first glance, a bit puzzling. What is happening is that our model is fit using transformations that are manually adjusted: we remove nonintuitive fluctuations and nonmonotonicities because we believe they reflect sample variation, not persistent relationships. This is where judgement helps a modeler, in that it is naïve and dangerous to simply 'let the data speak for itself', as invariably powerful statistical methods will find patterns and correlations that are not robust. We could have fit the data better, but didn't, because of *a priori* restrictions on how inputs relate to defaults. The cross-validation approach used no a priori restrictions (mainly within the transformation functions), and actually ended up fitting this set of data better than our final model. This should give the reader a sense of our focus on true out-of-sample performance, and our appreciation of the dangers of overfitting statistical models.[24]

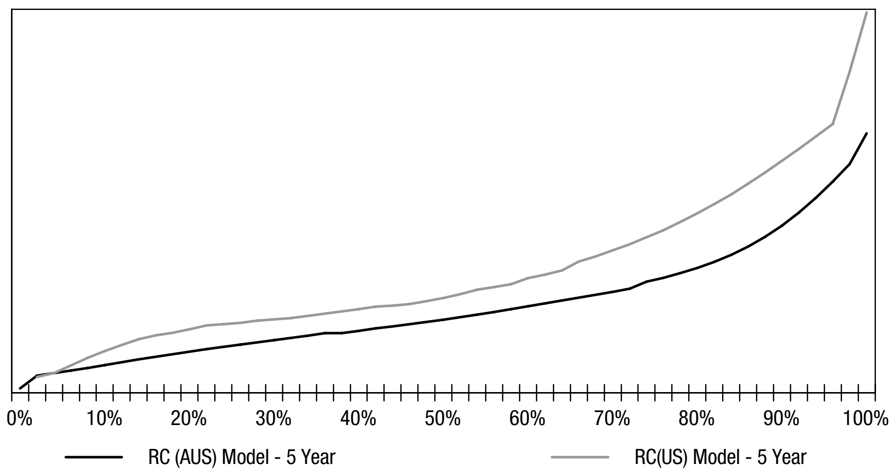
*23 Z-Score= 6.56*Working Capital/Assets+3.26*Retained Earnings/Assets+6.72*EBIT/Assets+1.05*Net Worth/Liabilities.*

24 Going one step further, we could have chosen highly nonlinear and interaction functional form that, in the limit, would be perfectly predictive for our sample. Clearly this would be a meaningless model.

While there appears relatively little difference in power by constructing a 'native' model, there is a large difference in DP. Examine **Figure 10** below. Note that the US and Australian models, as applied to Australian companies, generate significantly different DP distributions. The US model as applied to Australia is 'too pessimistic', as it generates DPs a full 1.3% higher on average than does the Australian model. This is the importance of calibration, of making sure that the model, when applied to a large number of credits, will produce aggregate numbers consistent with reasonable estimates of default probabilities from a 'top-down' perspective. Calibration, as opposed to power, is less robust across countries and suggests that it is more important to recalibrate a model for local conditions, as opposed to re-weighting the input.

Figure 10

DP Comparison Between Models on Australian Model

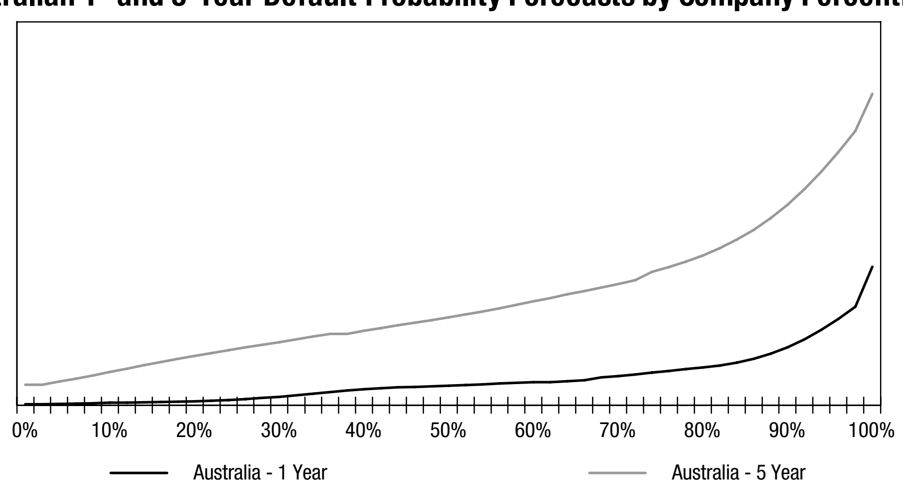


The US model as applied to Australia generates significantly higher DPs

A look at the DPs generated by RiskCalc, by percentile, helps understand how different forecast horizons imply different usages. In **Figure 11** note that for the 1-year horizon, the DPs are basically 0 until starting to rise at about the 75 percentile. Thus for short horizons, the model is most useful only at examining the shades of gray amongst the worst 25% of companies; among the best 75% distinctions are without much economic significance. For the longer horizon the inflection point is again around the 75th percentile, but prior to this point there is a more meaningful distinction among the better credits. Thus the longer horizon is better at distinguishing credit quality among the better credits, and in both cases the model works best at distinguishing between the bad and the very bad, as opposed to the good and the very good. Clearly as a monitoring tool it would be extremely valuable in deciding which credits to reexamine, and in deciding which credits are deserving of expedited review.

Figure 11

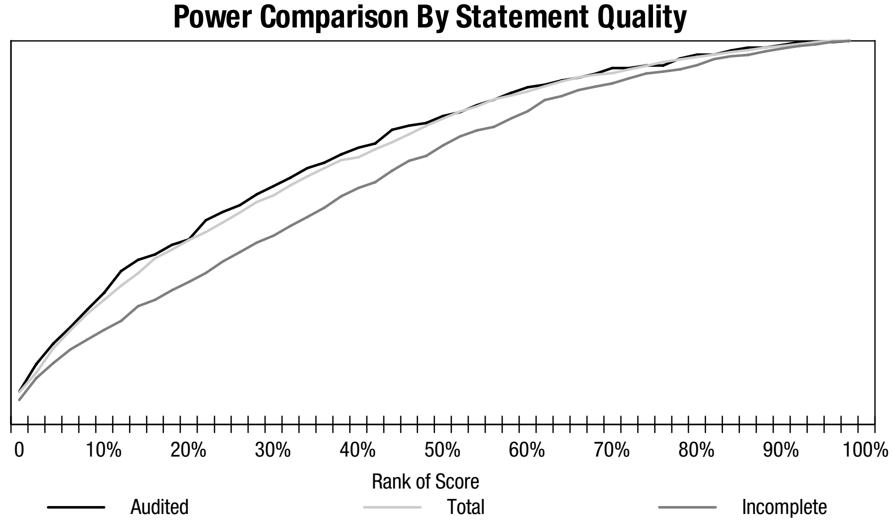
Australian 1- and 5-Year Default Probability Forecasts by Company Percentile



The 75th percentile is the inflection point for default probability prediction

Invariably RiskCalc is computed using unaudited or incomplete financial statements. It is important, therefore, that power be maintained across these subsamples, as otherwise the usefulness would be severely constrained. **Figure 12** below shows the relative performance of RiskCalc across these data subsamples for Australian firms. While performance does degrade slightly for incomplete statements, the model retains its power on these statements. More information is clearly better, but even with incomplete statements the model provides power.

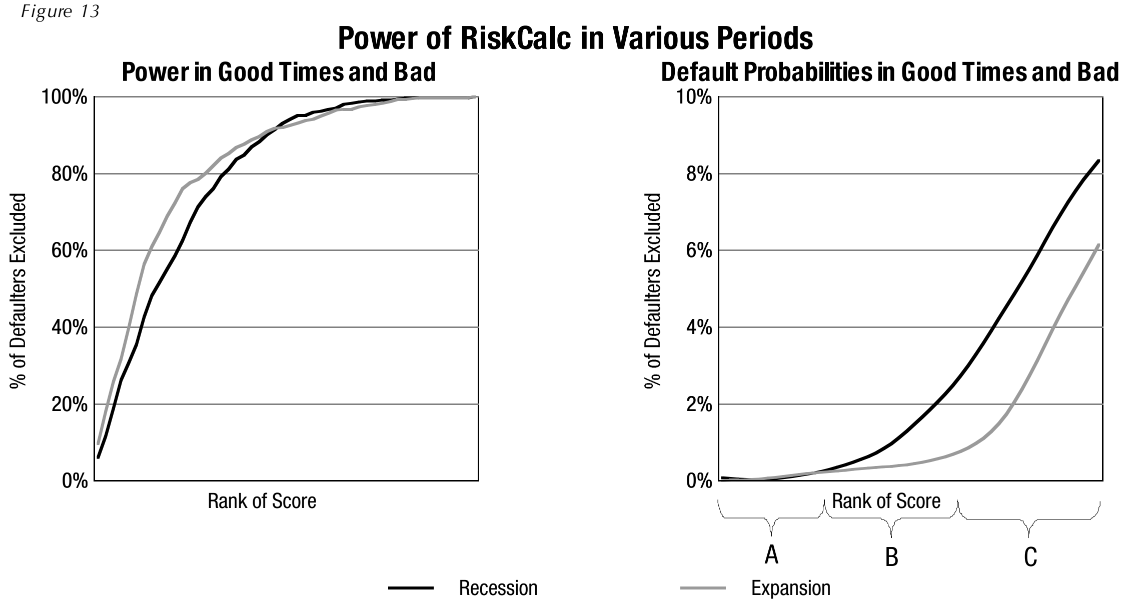
Figure 12



RiskCalc performs better for audited and complete statements

Figure 13 below illustrates the statistical power of RiskCalc in various periods. As was mentioned above, RiskCalc is estimated primarily during nonrecessionary years. In order to gauge its robustness during recessionary periods, we examine the power and DPs of the model during recessionary and nonrecessionary times for public companies.

Figure 13



Recessions both increase and broaden the base of defaulters, lowering the measured power of default model in these periods

Applied to US public companies, **Figure 13** shows that the power of the model decreased during the 1990 US recession. Recessions act like a shock to the system, affecting weaker firms with higher probability than they do stronger firms, and in bad times a firm's financial statements are slightly less powerful at explaining the relative default propensities. Looking at the chart in **Figure 13b**, consider regions A, B, and C. In bad times, the really bad firms (C) default, while moderate firms (B) and excellent firms (A) don't default. In bad times, however, section C still shows more defaults, but now section B is a gray area; a new class of firms that were previously almost never defaulting are now defaulting at low, but significant rates. This adds gray to a situation that was previously black and white, which the lower power of the model reflects. Mathematically, one can see why this happens-it's as if the shock to the default probability is additive so all models become less powerful. Thus power drops when the aggregate sample default probability rises. Most importantly, the model is still a powerful predictor.

Implementation Tips

There are two key counterintuitive aspects of the output that one should bear in mind when implementing RiskCalc. First, our model has a bias towards parsimony that can leave it somewhat counterintuitive. It ignores items we know to matter solely because it doesn't have data on them. For example, the model ignores industry differences, or various alternative views of profitability. This is only because we do not have enough data to sufficiently accommodate this information. It is a counterintuitive statistical fact that while adding information to a model will always help it better explain a given set of data, for prediction additional information worsens forecasts after a certain point. Like a good omelet, a good model doesn't contain everything, but only a subset of all that is available and individually good.

As Moody's receives more data we will increasingly nuance the model to account for more factors. Acknowledging this admits the imperfection of the current model, and makes it appear to some as inferior to models that seem to account for 'everything'. Yet building a model based on perfect data that doesn't exist is surely worse than building a model based on the best set of imperfect data available.

Another counterintuitive result of the model is that its implied DPs often appear 'too low'. Many users are used to higher default rate projections for individual loans. The average default probability for middle market loans of 1.7% per year seems to make perfect sense until one actually sees the individual credits to which this is applied, and one considers that this is consistent with Ba2 default probability (most people consider private credit in the B2-B1 range on average, not the Ba2 range). While the 1.7 % figure is documented, you should recognize that we want the model to be unbiased. That is, it represents our best statistical estimate of the future default probability. In contrast, a natural inclination of an underwriter is to be pessimistic, as the cost to being too optimistic is larger than that for being too pessimistic. In practice you may wish to adjust the results by multiplying the default probability by a constant such as 1.5, in order to better accommodate your prudent bias, too accommodate 'gaming', or simply because your old scale had a mean implied default probability of 3.0% and you feel the new scale should be modified slowly.[25]

To adopt the unadjusted output into pricing, incentive compensation, or strategic decisions is not simply politically naïve, but practically suboptimal. Not that we do not truly believe in the power and calibration of our model, just that prudent incrementalism is a good strategy in implementing any enterprise-wide performance measurement or pricing system. For example, if in practice a bank has used a severely miscalibrated DP for pricing models for many years, there must exist some counteradjustments in the relationship pricing schedule, otherwise there would have been an obvious problem (zero loan if priced too high, many bad loans if priced too low). Changing only the DP to the 'correct' calibration would be the opposite of an improvement from the bank perspective, since now the previously useful counteradjustments would cause the ultimate pricing schedule to be biased. It suggests an incremental approach is necessary for applying new DPs into performance measures and pricing schedules, while a more straightforward approach can be used for those uses which require only ordinal information, such as monitoring, and expediting underwriting.[26]

25 If the current system implicitly has a mean DP of 5% and the 'correct' DP is actually 2%, it is often prudent to move slowly toward the 'correct' DP, by perhaps using a DP of 4% as a first adjustment.
26 Users of Moody's Financial Analyst will be able to create customized mappings that relate DPs into internal grades.

Target For RiskCalc

Our target is the middle market universe, which we define as any company that is above $100,000 in assets yet does not have publicly traded equity. Thus a firm listed on the New York, Toronto or Sydney exchange, or even one with shares traded over-the-counter, is not optimally suited for this model because it ignores useful information: market value and its volatility.[27]

Two factors drive the lower bound of applicability. As companies become smaller the credit information of the main proprietor(s) becomes more important than the financial statements, which is often recorded in bureau scores, such as those offered by Equifax, Experian, or TransUnion. In some sense, a $50,000 loan to a company is more like a credit card exposure than a corporate bond. Secondly, the financial statements of these smaller firms tend to be both less accurate (fewer audited statements) and less timely, with delays of up to 6 months common. While at the lowest end of the spectrum it is best to focus solely upon the proprietor, there is a middle area where the proprietor credit rating and the firms financial statements are both important, and a a level above this at which the proprietor's credit information is not relevant. RiskCalc is clearly for the latter two areas, with the gray area being between $500K and $3MM in sales.

Conclusion

RiskCalc's Australian model provides a powerful, transparent and consistent model for evaluating middle market credits. Careful attention has been paid to how financial ratios differ between Australia and North America, how these ratios relate to default, and how data tend to exist in Australian databases. RiskCalc is by no means a sufficient measure of risk, but as an aggregator of financial statements into a validated, meaningful number that allows apples-to-apples comparison of portfolio risks, it is extremely useful.

27 Our model for that is Moody's RiskCalc for public companies.
28 This is mainly because smaller firms are more idiosyncratic, and hence less predictable, due to their lack of diversification. Also, smaller firms have noisier (less accurate) financials, which adversely affects the forecast power.

Appendix A: Testing Metrics

Power Curves

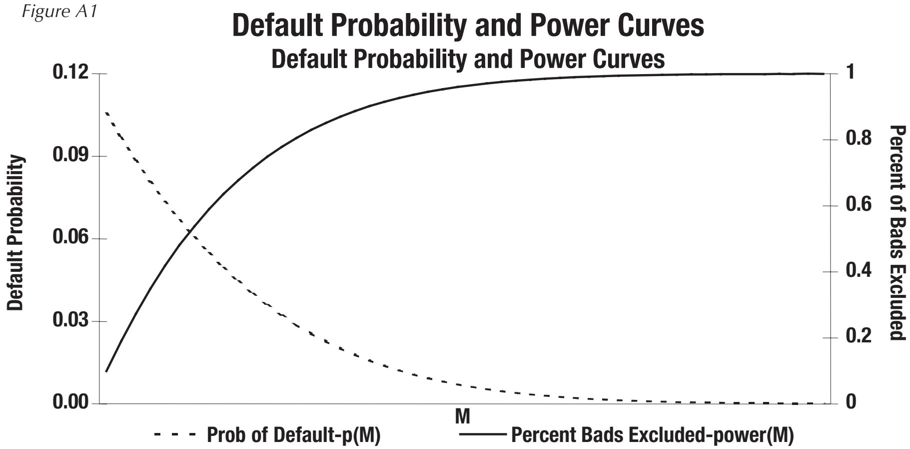
A power curve[29] is constructed by plotting, for each threshold, the proportion of defaults excluded at various levels of sample exclusion. The vertical axis measures the percent of defaults excluded conditional upon excluding various percentage levels of the sample. Thus if using a score to exclude 50% of the sample caused it to lose 80% of the defaulting companies, the power curve would go through a line corresponding to x=0.5 and y=0.8. Accuracy is indexed more precisely by the amount of area under the curve, which increases as the curve bends.

$$power\ (b) = \frac{\Sigma_{t=1}^{b} p(t)}{\Sigma_{t=1}^{B} p(t)} = \frac{defaults\ excluded\ at\ b}{total\ defaults} \qquad (0.4)$$

Here B is the total number of bins (often 10 for illustrative purposes), and b is a particular bin, . The power at bin b represents the sum of all the defaults in the 'worst' fraction b/B of the scores, as ranked by the metric M.[30]

The net result is Figure A1 below, which shows the probability of default for a level of M, and statistical power, which pertains to the nature of all the data up to a level of M. In this case we rank-order the firms from risky (left) to less risky (right), so that the P(M) and Power(M) correspond. The graph shows a particular case. This type of model would quickly excluded most of the bad companies: a 20% exclusion of the worst companies according to the M score would exclude 70% of the future defaulters.

Figure A1



There is a one-to-one correspondence of power and probability of default by rank-order, in that for any point t along a default metric:

$$p(t)= \overline{p} \ * \ \frac{\partial\ power(t)}{\partial t} \qquad (0.5)$$

where $\overline{p}$ is the mean default probability.

While the graphical or tabular display of power is informative, and has the advantage of allowing one to examine power at variety of thresholds, it is useful to aggregate the power curve information into a single number that allow unambiguous comparison. One such measure is the area under the power curve. A model more bowed out towards the left will have a greater area, and be more powerful on average.[31] Using the area under the power curve implies that there can exist threshold levels such that a model with a smaller total area has a momentary advantage. Thus area is not a measure of global or complete dominance, just an intuitive measure of dominance on average. The area can be calculated using equation (0.4) above, specifically

$$Area= \frac{1}{B}\ \Sigma_{b=1}^{B}\ power(b) \qquad (0.6)$$

29 Also known as Gini curve, CAP plot, Lorentz curve, ordinal dominance graph, or ROC curve.
30 Since defaults excluded 'at bin b' is ambiguous-it could mean 'up to bin b' or 'up to and including bin b'-we calculate the area using the average of the two methods. In practice this appropriate adjustment is not significant.
31 Using the total area under the powercurve implies that there can exist points, or threshold levels, such that a model with a smaller area has an advantage at a particular threshold level. Thus it is not a measure of global or complete dominance, just an intuitive measure of dominance on average.

where B is the total number of bins. If the Area is greater for one model than another, it is more powerful.

Of course to really compare models you not only need an aggregate measure of power, but a standard error on this metric. The standard error of this area is approximated well by the following formula:

$$\sigma_{Area} = \sqrt{\frac{1}{BD} \sum_{b=1}^{B} p(b)(1-p(b))} \qquad (0.7)$$

Appendix B: Power Curve Construction Details

The testing approach is as follows. We take each defaulting firm, and find the score 12 months prior to the default date. If a score does not exist on this date, we move backward in time until we reach 24 months prior to the default date. If no score is present we exclude the defaulting observation from this particular test. We find the score's relative rank among all scored firms (defaulters and nondefaulters) at the time of the score. This is done by using the calendar year from which the score was taken. While this explicitly corrects for business cycle, in that aggregate fluctuations in scores are anticipated, for models that use only financial statements this consideration is of little practical importance.[32] Each defaulter is then mapped into a percentile, and this collection of percentiles is the basis by which the power curve is created. Specifically, given a collection of percentiles of defaulting firms $\{\phi j\}^J_{j=1}$, where J is the total number of defaulting firms, the power for each bin is simply

$$power\ (b) = \frac{1}{J}\ \Sigma^J_{j=1}\left\{1 \middle| \phi_j \leq \frac{b}{B}\right\} \qquad (0.8)$$

Where $\left\{1 \middle| \phi_j \leq \frac{b}{B}\right\}$ is an indicator function equaling 1 if defaulting firm j was in a percentile of scores lower than $^b/_B$, K months prior to the default date. As mentioned above, we take the score in the month closest to K up to a limit, so in effect the percentile is from K to K+ΔT back from the default date. For example, for a 1 year default probability tests we would take a default on 5/98, and move back to 5/97 to find the percentile of the RiskCalc score in 1997 using that month. As is most probable, the statement date is not exactly at 5/97, and so we must go back in time, to 4/97, then 3/97, etc., until we find the date at which we have financial statements. For the short term test, we go backward in time starting 90 days prior to default and moving backward in time to two years prior to default; for the long term test we go backward in time starting 5 years prior to the default date and move forward in time until we get to 36 months prior to the default date.

By going backwards at least 90 days from the default date, we avoid the misleading results that come from model performance over irrelevant time periods, such as 60 days after a statement date. Predicting default of very short horizons, such as less than 90 days, is basically useless, as very few statements are completed within this time. Many lenders take 6 months to be confident that most of their middle market exposures have delivered their latest annual statements.

By using defaulting firms *once* in the creation of a set of percentiles of defaulted firm scores, we avoid double counting firms, which can also cause problems, especially with standard errors that usually assume independence within the sample. The short run horizon can than be easily extended to the longer run horizon where double counting is especially problematic, in that a 5 year default probability would double count 'defaulting observations' if one simply used statement dates and firms and measured which firm-year observations had defaults within 5 years. For the longer run default curves we take default date and move backward 60 months to find a score (which implies a financial statement. If one doesn't exist, we move forward in time until a score is found (up to 36 months within the default date). The result is a collection of percentiles which measures the relative ranking of defaulting firms among all companies 3-5 years prior to default.

---

32 *Financial ratios are not very cyclical.*

References

Falkenstein, Eric, 2000, *Validating Commercial Risk Grade Mappings: Why and How*, *Journal of Lending and Credit Risk Management*, February.

Falkenstein, Eric, Carty, L.V., and A.K. Boral, 2000, *RiskCalc for Private Companies: Moody's Default Model*, Moody's Invetors Service Special Comment, May.

Gupton, Greg, S.C., Gates, Daniel, and L. Carty, 2000, *Bank Loan Loss Given Default*, Moody's Investors Serive Special Comment, November.

Herrity, J., Keenan, S.C., Sobehart, J.R., and L. Carty, *1999*, *Measuring Private Firm Default Risk*, Moody's Investors Service Special Comment, June.

Sobehart, J.R., S.C., Keenan, and R. Stein, 2000, *Benchmarkeing Quantitative Default Risk Models: A Validation Methodology*, Moody's Investors Service Special Comment, March.

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

Report Number: 62265