A Practical Review and Test of Default Prediction Models

by Jorge R. Sobehart and Sean C. Keenan

The authors, members of Citigroup’s Risk Architecture group, examine the benefits and limitations of two alternative classes of models for default risk assessment. The contingent claims analysis is a structural approach based on an option-theoretic view of a firm’s equity and liabilities. But even CCA models often show poor power in predicting default events in precisely those cases where these models should hold strongest. By introducing additional credit information, hybrid models are able to overcome some of these limitations and increase model performance in the critical near-default region.

Competitive, legal, and regulatory pressures continuously demand more sophisticated approaches to managing credit risk in commercial lending and investment. Quantitative credit risk models that detect a broader spectrum of stressful conditions are key to the monitoring and managing functions required for active portfolio risk management.

While such models can offer more timely, wider, and more cost-effective coverage than teams of professional analysts, their performance characteristics are not well known. Various commercial vendors now offer competitive products with claims of precise anticipation of default events and major changes in value. Recently, tests of portfolio management rules based on a commercial product performed by J.P. Morgan Credit Strategy suggest that some of these claims are incorrect or exaggerated.1

First There Was Altman, Then There Was Merton

By the mid-1980s, there were two main approaches to quantitative default risk modeling: the statistical approach pioneered by Edward I. Altman in 1968 and the...
contingent claims analysis (CCA) approach of Robert C. Merton and others in 1974. The former, based on fundamental credit information, uses econometric techniques to determine the relationship between the event of default on obligations and market information, accounting variables, and credit opinions, such as ratings. CCA is a structural approach based on an option-theoretic view of the firm’s equity and liabilities. Merton showed that a firm’s equity could be viewed as a call option on the firm with strike price equal to the face value of its debt. That is, a firm’s default risk could be modeled with the same approach used (by Fischer Black and Myron C. Scholes in 1973) to price an option using stock prices and a few key parameters taken from the firm’s balance sheet.

Unfortunately, even models with sound theoretical underpinnings, such as CCA, often show poor power in predicting default events in precisely those cases where these models should hold strongest. CCA models also are limited in performance during periods of volatile markets (providing too many sell signals or false positives for default prediction). In addition, they are unreliable in predicting credit spreads and bond prices. This is largely the result of simplifications and abstractions required to make the CCA models tractable, but which are unrealistic and tend to seriously degrade model performance in practice.

And Then Came the Hybrids

By the late 1990s, a third approach had emerged that acknowledged not only the value of timely information like stock prices and volatility, but also the importance of fundamental credit information. This third approach led to the introduction of “hybrid” models [see Footnote 1 at the end of this article]. The key idea underlying hybrid models is that default is a contextual event.

No credit officer would ever extend credit solely on the basis of a strong stock price or low stock volatility, nor would credit extension be ruled out by a sudden drop in the market value of equity as implied by idealized CCA models. Hybrid models show that, while market equity information can be extremely valuable, it is most useful when it is coupled with an understanding and explicit consideration of the fundamental information of the firm and its business environment. A detailed examination of a firm’s balance sheet, income statement, and cash flows remains a critical component of credit risk assessment. Finally, because the additional accounting and credit information is not used to forecast equity prices but to enhance the definition of the event of default, hybrid models do not contradict any hypothesis regarding market efficiency or the no-arbitrage conditions required in CCA models.

Now that commercial applications of hybrid models have become available, it is important to contrast their performance with the market-dominant CCA models. This requires answering two questions:

1. How accurately do CCA models describe actual default experience?
2. How well does this compare with the hybrid models?
Performance Testing

Alternative hybrid models have drawn the attention of financial institutions and commercial vendors for good reason: one study’s conclusion that they could have provided early warning signals for such high-visibility cases as Xerox, Owens Corning, Edison International, and PG&E, and many others. This study, by Sobehart et al., found that additional information could improve the predictive power of a Merton model, providing a more robust model of credit quality.

A second study, by Kealhofer and Kurbat, actively criticizes the hybrid nature of the model, concluding that agency ratings and additional market and financial information do not add any power to options-pricing based models, findings that squarely contradict those of the first study. Despite their claims, however, their empirical tests do not actually compare any version of the Merton model to a hybrid model, but compare their own Merton model (KMV implementation) to a handful of accounting ratios, Moody’s ratings, and a simple Merton model reported in the literature.

A review of the empirical tests has indicated that the methodology and conclusions of Kealhofer and Kurbat are biased as a result of several factors, including:

- Incorrect specification of the hypothesis to be tested.
- Incorrect inference about the performance of hybrid models based on an unrelated model.

This article compares their version of the Merton model to the published hybrid model using U.S. nonfinancial firms, including both rated and not-rated obligors. Both models produce estimated default probabilities over a one-year time horizon, known as expected default frequency (EDF) and expected default probability (EDP), respectively. The study’s data set:

- Includes more than 27,000 firm-year pairs of annual observations for both models for the period December 1995 to December 1999. Each pair of observations measures the obligors’ credit quality at the beginning of the following year.
- Contains 349 default events occurring during the period January 1996 to December 2000. The numbers of firm-year observations and defaults in each specific year are listed below.

<table>
<thead>
<tr>
<th>Year</th>
<th>Firm-years</th>
<th>Defaults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>5,099</td>
<td>41</td>
</tr>
<tr>
<td>1997</td>
<td>5,361</td>
<td>50</td>
</tr>
<tr>
<td>1998</td>
<td>5,713</td>
<td>95</td>
</tr>
<tr>
<td>1999</td>
<td>5,755</td>
<td>101</td>
</tr>
<tr>
<td>2000</td>
<td>5,487</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td>27,415</td>
<td>349</td>
</tr>
</tbody>
</table>

The observations were ordered by model output—from riskier to safer—and the defaults used were those that occurred within the following 12 months of each annual observation. Figure 1 shows the cumulative fraction of defaults as a function of the ordered population of obligors. Interpretation of this plot is straightforward; the higher and steeper the curve the better the model is at differentiating between defaulters and healthy firms.

Figure 2 shows similar accuracy curves for individual years, indicating that the hybrid model outperforms the equity-based model consistently. Figures 1 and 2 suggest that hybrid models improve over state-of-the-art commercial applications of options-pricing models, reducing the number of false positive signals (or sell signals).

From the practitioner’s point of view, Figure 2 tells a very compelling story. More precisely, it shows that the misspecification of CCA models is more apparent for high-risk firms and becomes most severe as we approach the very point these models were designed to identify. By introducing additional credit information specifically to enhance the identification of the case-specific default point, hybrid models are able to overcome some of these limitations and can increase model performance in the critical near-default region. The result is a more reliable tool for analyzing the risk of firms in the most profitable segment of credit quality.

Why Hybrid Models Should Work Better

In the standard CCA framework ruled by idealized efficient markets, credit assessment simply reduces to identifying changes in stock prices and volatility as a reliable source of credit quality information. This provides a short-sighted understanding of credit quality based on unrealistic assumptions about the actual dynamics of the debt markets and the role of arbitrage. In the real
world of lending, value uncertainty and potential arbitrage situations are rampant because most institutions’ assets and liabilities do not possess the same idealized characteristics and liquidity required by CCA models. For example, unlike true option sales, the issuance of debt is not associated with any bona fide hedging activity that would tend to result in a high volume of incremental reallocative trades. Practically speaking, large denominations and thin trading in the bond and loan markets, against a backdrop of fluctuating stock prices, are inconsistent with the idealized conception of debtholders as writers of perfectly liquid options on the unobservable assets of the firm proposed in pure CCA models.

Another key problem with CCA models based on market equity stems from the fact that even fully informed equity prices are marginal prices and, therefore, primarily reflect marginal reallocative supply and demand conditions rather than the value of the aggregate capital stock of the firm.

Differing views on the best deployment of the firm’s assets can create a “control premium” and a steep supply curve for equity shares. The huge price swings that accompany takeover bids prove this point. Tightly controlled firms with inefficiently deployed assets may trade at a discount. The opposite situation, where equity trades at a hefty premium, can also occur as a result of speculative valuation, such as in the dot-com and telecom frenzies. It’s unknown how much these supply/demand effects may be affecting market prices, so they are a potential source of uncertainty for equity valuation.\footnote{This issue is fundamental in the credit risk context because commercial applications of CCA models estimate the implied market value of the firm’s assets and volatility from stock prices, which...}

Figure 1

![Figure 1](image1.png)

Power curves compare the overall default predictive performance of any default prediction model. The higher and steeper the curve, the better the model is at differentiating between defaulters and healthy firms.

Figure 2

![Figure 2](image2.png)

Accuracy comparison of a hybrid model (solid line) and an equity-based model (dashed line) for the sample described in Figure 1 and the table seen on page 56. The higher and steeper the curve, the better the model is at differentiating between defaulters and healthy firms.
are then used to estimate probabilities of default or credit spreads. Therefore, it is important to understand how the uncertainty of equity and debt markets can bias these estimates of credit quality. For example, a market correction might reflect a change in investors’ preferences that could generate a sudden decrease in the firm’s stock price without necessarily implying an increase in the probability of default of the firm, as models derived within the CCA framework would indicate.

Anyone familiar with commercial applications of the CCA framework (such as the models in Figure 1) would notice that the swings in stock volatility over the last couple of years were not followed by a surge in defaults to the extent predicted by these models. Furthermore, equity-based models frequently assign speculative grade ratings (probabilities of default) to well-known investment grade corporations and financial institutions. This effect is exacerbated during periods characterized by stock market volatility. During these periods it is not unusual for equity-based models to assign a low credit score to a large fraction of the population of firms, generating undesirable sell signals or “false positive” default events as shown in Figure 2.

A hypothetical example, based on the standard Merton model, illustrates this point. An additional source of equity volatility was included to simulate the uncertainty of investors. Then, the European call option formula from the Merton model was inverted to obtain implied values of the firm’s assets and volatility as a function of the true assets and volatility, and the par value of debt (the “exercise price” of the option).

Finally, the probability of default was calculated using the true assets and volatility of the firm, and the values implied from equity with the additional source of randomness. Figure 3 shows the probability of default— for a one-year time horizon— as a function of the firm’s effective leverage obtained with the true value A and volatility of the firm’s assets and the face value of debt D. Figure 3 also shows the probability of default as a function of the leverage estimated with the assets value and volatility “implied” from equity using the Merton model with an additional source of noise. Note that for highly leveraged firms (point A in Figure 3) the probabilities of defaults for true and implied values are similar. However, for highly capitalized firms (point C) the probability of default obtained by using implied assets and volatility assigns a lower credit quality (point B) to these firms than it should. Thus, pure equity-based models of risky debt can easily misclassify investment-grade firms as speculative-grade firms when markets are volatile. This is a consequence of the additional uncertainty introduced by the market participants, which is erroneously interpreted as asset volatility by contingent claims models. These models will always make these errors because they are based on the assumption that market participants are perfectly informed, and stock prices can only reflect “true” changes in the value of the firm’s underlying assets rather than marginal reallocative supply and demand conditions. Because hybrid models provide a more comprehensive credit assessment that combines market information and accounting information, they help reduce the impact of valuation uncertainty and model bias discussed above, as evidenced by their superior default predictive performance as shown in Figures 1 and 2.

The Bottom Line

It seems natural that information widely used by credit analysts and loan officers should play a role in a quantitative default prediction model. Such information can improve the performance of contin-
gent claims analysis models. But can this more traditional credit information be brought to bear on the practical monitoring and managing of risky credits without giving up the advantages in speed and coverage of a quantitative model? The answer is a resounding yes. Hybrid models that do make use of such information outperform their market-based counterparts in terms of identifying future defaulters and avoiding “false positive” signals. They do this without sacrificing speed or coverage. The additional inputs they require are more easily and cheaply obtained than those needed for the CCA part of the model. Finally, because these models can also be used to assist institutions in the determination of capital requirements, even a marginal improvement over an existing model can generate huge capital savings and an immediate improvement in an institution’s risk-adjusted return on capital.

Contact the authors by e-mail at jorge.sobehart@citicorp.com and sean.keenan@citicorp.com

Footnotes

Other References