The Performance of Structural Models in Pricing Credit Spreads

Manuel Rodrigues
(Cranfield University School of Management)

Vineet Agarwal*
(Cranfield University School of Management)

Version: 15 January 2012

*Corresponding author:
Vineet Agarwal
School of Management
Cranfield University
Cranfield
MK43 0AL

Tel: + 44 (0) 1234 751122
Fax: + 44 (0) 1234 752554
E-mail: Vineet.agarwal@cranfield.ac.uk

ABSTRACT

There is a lack of consensus in the literature regarding the performance of structural models even though they have been used as a standard in credit risk modelling for the last thirty years. This paper tests the performance of three structural models, Merton (1974), Collin-Dufresne and Goldstein (2001) stationary leverage ratio model and CreditGrades™ (2002) against the credit default swap spread using much larger dataset of 320 corporates over a longer time period than the existing literature.

We find that, consistent with existing literature, while the structural models have significant explanatory power for the cross-section of CDS spreads, all three models underpredict the observed credit spreads, and the underprediction is not driven by any ratings category or sector. We also show that while the three models produce spreads that are highly correlated with each other, they carry information incremental to each other. However, our evidence demonstrates that other firm specific variables can explain a significant part of observed credit spreads indicating that the observed underpricing is driven by actual spreads reflecting additional information such as liquidity and residual credit risk. In fact, the actual spreads are lower than implied spreads once additional credit risk and CDS liquidity information is taken into consideration.

Keywords: Credit spread pricing, Bid-offer spreads, Credit ratings, Structural model

JEL: G12, G13, G24

1. Introduction

Corporate bonds are priced by charging a premium over the corresponding government bond yield (risk-free-yield). This premium is determined by the creditworthiness of the corporate borrower. Commonly called "credit spread," this gap represents the extent to which the company is at risk of defaulting on the bond, and compensates investors for bearing the risk.

The first structural model of Merton (1974) for pricing corporate bonds assumes that a firm defaults if the value of its assets is below its outstanding debt at the time of servicing the debt. Although this model has been widely adopted for valuing corporate bonds, the underlying assumptions of the original model often do not reflect economic reality.[1] Several new structural models have refined the original framework of Merton (1974) by relaxing one or more unrealistic assumptions. For instance, Black and Cox (1976) introduced the first passage model whereby default could occur before maturity. Geske (1979) considered the risky coupon bond as a compound option and defined the default barrier as the market value of remaining debts. Longstaff and Schwartz (1995) modelled stochastic interest rates following a mean reverting process, where coupon bonds were considered a weighted sum of zero coupon bonds. Leland (1994), and Leland and Toft (1996) assume that firms could issue equity to service debt and that default occurred when the value of equity goes to zero. Finally, Collin-Dufresne and Goldstein (2001) drop the assumption of a constant default boundary, and model leverage as a stationary process.

However, in spite of the growing number of structural models, there has been very limited empirical testing in the literature. Jones, Mason and Rosenfeld (1984) is among the first studies to test the predictive ability of the Merton (1974) model. Using data on 27 US firms with simple capital structures between 1975 and 1981, they find that the observed bond prices are much higher than those predicted by the model. Ogden (1987) reaches same conclusion using a sample of 57 new bond issues by US industrial corporations between 1973 and 1985. Lyden and Saraniti (2001) study a sample of 56 US corporates with a single outstanding bullet bond between 1988 and 1999. They also conclude that both, Merton (1974) as well as Longstaff and Schwartz (1995) models underpredict the observed bond spreads.

[1] The model assumes, *inter alia*, default can only occur at maturity, interest rate is constant over the life of the option, default is never a surprise, the price of the underlying stock one period ahead follows a log-normal distribution, trading in assets takes place continuously, and there are no transaction costs or taxes.

However, Eom, Helwege and Huang (2004) find contrary evidence by implementing five structural models on a sample of 182 bonds between 1986 and 1997.[2] They find that while Merton (1974) and Longstaff and Schwartz (1995) models do on average, predict spreads that are too low, the other three models produce estimates that are higher than those observed. However, they find that the absolute prediction errors of all the models are quite high, indicating poor predictive ability. Similarly, Delianedes and Geske (2001) look at the credit spreads of 470 bonds between 1991 and 1998, and find that Merton (1974) model estimates are able to explain only about 5% of the observed spread for AAA rated bonds and 22% for those rated BBB. They argue that bond spreads are mainly explained by market returns, market volatility, and bond liquidity.

Ericsson, Reneby and Wang (2007) is the first study that uses credit default swaps in addition to bonds to test the empirical performance of structural models. Using a sample of 23 industrial bonds between 1997 and 2003, and the same five structural models as in Eom et al. (2004), they find no evidence of underprediction of CDS spreads. They also find that the results using bond spreads are sensitive to how the credit spread is measured. Using the spread over treasury, they find, consistent with existing literature, underprediction of bond spreads, however, there is no such evidence if the spreads are measured over the swap curve. Further, using CDS spreads, they find no evidence of role of liquidity or credit ratings in the residuals.

In this paper we test the performance of three structural models in pricing credit default swaps. We select a generalization of Merton's model (1974), since it is the seminal credit spread model. Collin-Dufresne and Goldstein's (2001) stationary leverage ratio model is chosen given its apparent superior performance according to the literature (e.g. Eom at al., 2004; Huang and Zhou, 2008), and Creditgrades (2002) as it is the current standard model used by a wide range of credit market practitioners.

We extend the existing literature in two ways. First, we provide a much cleaner test of the predictive ability of structural models by using CDS spreads rather than bond spreads used in the existing literature (with the exception of Ericsson et al., 2007). CDS spreads arguably provide a cleaner proxy for default risk because they are more liquid than corporate bonds and are not influenced by parameters such as call provisions or coupons. Secondly, our longer time period of 7 years that includes the current credit crisis period, and larger sample size of 320 corporates allows much stronger inferences. While we find

[2] The models are Merton (1974), Geske (1977), Longstaff and Schwartz (1995), Leland and Toft (1996), and Collin-Dufresne and Goldstein (2001).

underpricing of credit spreads as in existing literature, we also provide evidence for the first time that this is driven by CDS spreads reflecting liquidity premium as well as residual credit risk information from credit ratings. Once these factors are taken into consideration, we show that actual CDS spreads are actually lower than the implied spreads. We also show our findings are not influenced by the current credit crisis.

The rest of this paper is organized as follows: Section 2 presents the theory behind the Merton, Collin-Dufresne and Goldstein, and Creditgrades models; section 3 focuses on the models' implementation, explains the data and methodology used; section 4 discusses the empirical results; and section 5 concludes.

2. Theory and model description

We test three market-based structural models, specifically Merton (M), Collin-Dufresne and Goldstein (CDG), and CreditGrades (CG).

2.1. Generalization of Merton (1974)[3]

The Merton (1974) model assumes that a firm with an initial value of V_0 issues a zero-coupon bond with a face value of D and a maturity of T. The critical assumptions behind the model are that markets are perfect, that trading in assets takes place continuously, and that the firm's value follows a geometric Brownian motion under a risk-neutral measure:

$$\frac{dV_t}{V_t} = (r-\delta)\,dt + \sigma\, dz^Q(t) \tag{1}$$

where:

r = the risk free rate,

δ = the dividend yield,

σ = the volatility of the firm value, and

$z^Q(t)$ = a Wiener process with a (r-δ) drift.

Default occurs at time T if $V_T < D$. Bankruptcy costs are assumed to be a constant proportion α of remaining firm value. The risky zero-coupon bond price is then given by:

[3] The difference between the Merton (1974) model and its generalization is that the latter relaxes the no dividends assumption.

$$P_M = D \cdot e^{-rT}\left(N\left(d^{Q}_{(\Gamma,T)}\right) + (1-\alpha)\cdot \Gamma \cdot e^{(r-\delta)T} \cdot .N\left(-d^{R}_{(\Gamma,T)}\right)\right) \tag{2}$$

Where:

$\Gamma = V_0 / D$ is the inverse-leverage ratio,

$N(\bullet)$ = univariate standard normal probability density function,

$$d^{Q}_{(x,T)} \equiv \frac{\ln x + \left(r - \delta - \frac{\sigma^2}{2}\right)\cdot T}{\sigma\sqrt{T}},$$

$$d^{R}_{(x,T)} \equiv \frac{\ln x + \left(r - \delta + \frac{\sigma^2}{2}\right)\cdot T}{\sigma\sqrt{T}},$$

T = maturity in years, and
∝ = bankrupcy costs

2.2. Collin-Dufresne Goldstein (2001)

This is a stationary leverage ratio model which assumes that firms tend to adopt a long-term stationary leverage ratio, i.e., they tend to issue debt when their leverage ratio falls below some target, and are more hesitant to replace maturing debt when their leverage is above that target. The model incorporates the idea that firms may adjust their outstanding debt levels in response to changes in firm value, which makes the stochastic leverage ratio revert to a mean value of $\bar{l}^{Q}$ in a constant interest rate framework.

This approach follows the following firm-value dynamics:

$$\frac{dV_t}{V_t} = (r-\delta)\,dt + \sigma\, dz^{Q}_{t} \tag{5}$$

The dynamics of other relevant variables under the risk-neutral measure are:

$$dy_t = \left(r - \delta - \frac{\sigma^2}{2}\right)dt + \sigma\, dz^{Q}_{t} \tag{6}$$

$$dk_t = \lambda(y_t - \nu - k_t)\,dt \tag{7}$$

where:

λ = the speed of adjustment to a company's target leverage,[4] and

ν = the leverage reversion level.

Equation (7) shows that k_t is mean-reverting and its target level is given by $(y_t - \nu)$. When k is less than $(y_t - \nu)$, the firm increases the outstanding book value of debt, and vice versa. Further, the log-leverage ratio is mean-reverting, and hence stationary, with a mean-reversion level of $\bar{l}^Q$:

$$dl_t = dk - dy = \lambda\left(\bar{l}^Q - l_t\right)dt - \sigma\, dz_t^Q \tag{8}$$

$$\bar{l}^Q = \frac{-r + \delta + \frac{\sigma^2}{2}}{\lambda} - \nu \tag{9}$$

Where τ is defined as the random time at which $l(t)$ reaches zero for the first time, triggering default.

Assuming that a risky discount bond pays a fraction of $(1 - \omega)$ of the T-maturity risk-free bond at the time of default, its price can be written as:

$$P^T(l_0) = e^{-rT}\left(1 - \omega \cdot Q(l_0, T)\right) \tag{10}$$

Where $Q(l_0, T)$ is the risk-neutral probability that default occurs before time T, given that the leverage ratio is l_0 at time 0. Discretizing time into n equal intervals, and defining date $t_j = \frac{j \cdot T}{n} \equiv j \cdot \Delta T$ for $j \in (1, 2, ..., n)$:

$$Q(l_0, t_j) = \sum_{i=1}^{j} q_i \qquad j=2,3,\ldots,n \tag{11}$$

$$q_1 = \frac{N(a_1)}{N\left(b_{(1/2)}\right)} \tag{12}$$

$$q_i = \left(\frac{1}{N\left(b_{(1/2)}\right)}\right) \cdot \left[N(a_i) - \sum_{j=1}^{i-1} q_j \cdot N\left(b_{i-j+\frac{1}{2}}\right)\right] \quad i = 2,3,..., n \tag{13}$$

[4] Which in this model is a mean reverting coefficient

$$a_i = \frac{M(i \cdot \Delta t)}{S(i \cdot \Delta t)} \tag{14}$$

$$b_i = \frac{L(i \cdot \Delta t)}{S(i \cdot \Delta t)} \tag{15}$$

$$M(t) = l_0 \cdot e^{-\lambda \cdot t} + \bar{l}^{Q} \cdot \left(1 - e^{-\lambda \cdot t}\right) \tag{16}$$

$$L(t) = \bar{l}^{Q} \cdot \left(1 - e^{-\lambda \cdot t}\right) \tag{17}$$

$$S^2(t) = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2 \cdot \lambda \cdot t}\right) \tag{18}$$

The term $Q(l_0,T)$ is the limit of $Q(l_0, T, n)$ as $n \to \infty$ .The q_i terms in equation (13) are defined recursively, which makes it straightforward to program this valuation expression and to calculate risky discount bond prices. Although $Q(l_0.T)$ is defined as the limit of $Q(l_0, T, n)$ according to Longstaff and Schwartz (1995), the convergence is rapid; and numerical simulations show that setting n=200 results in values of $Q(l_0,T)$ and $Q(l_0, T,n)$ that are virtually indistinguishable.

2.3. Creditgrades™

Creditgrades™[5] (CG) belongs to the class of structural models derived from the Merton model. Similar to the Merton model, the CG model assumes that the firm's value evolves according to a geometrical Brownian process. Default is defined as the point at which the value of the firm is below the default barrier, i.e., when it will not be able to meet its financial obligations. The default barrier is defined as the amount of firm assets that remains should the firm default. The average recovery rate that the debt holders receive upon default is L*D, where L is the average recovery of the debt and D is the firm's debt-per-share. The global recovery rate is assumed to follow a lognormal distribution with an L mean and the standard deviation of λ. The standard deviation of the global recovery rate is a way of modelling the uncertainties that affect the proper default barrier level. This is one of the prominent improvements claimed by the CreditGrades™ model over the Merton model.

[5] For details about the model, see Finkelstein et al. (2002)

The formula to determine a firm's survival probability up to time t is

$$P(t) = \Phi\left(-\frac{A_t}{2} + \frac{\log(d)}{A_t}\right) - d \cdot \Phi\left(-\frac{A_t}{2} - \frac{\log(d)}{A_t}\right) \tag{19}$$

Where

$d = \frac{V_0 \cdot e^{\lambda^2}}{L \cdot D}$ and $A_t^2 = \sigma^2 t + \lambda^2$

The spread of the CDS with maturity T, c^* can be calculated as

$$c^* = r(1-R) \cdot \frac{1 - P(0) + e^{r\xi} \cdot (G(t+\xi) - G(\xi))}{P(0) - P(t) \cdot e^{-rt} - e^{r\xi}(G(t+\xi) - G(\xi))} \tag{20}$$

Where $\xi = \lambda^2 / \sigma^2$ and the function G is given by

$$G(u) = d^{z+\frac{1}{2}} \cdot \Phi\left(-\frac{\log(d)}{\sigma\sqrt{u}} - z\sigma\sqrt{u}\right) + d^{-z+\frac{1}{2}} \cdot \Phi\left(-\frac{\log(d)}{\sigma\sqrt{u}} + z\sigma\sqrt{u}\right) \tag{21}$$

With $z = \sqrt{\frac{1}{4} + \frac{2r}{\sigma^2}}$

Where Φ is the standard normal cumulative distribution function, σ is the standard deviation of the firm's value.

The yield to maturity of a zero coupon bond can be obtained using

$$P^T(l_0) = e^{-Y^T \cdot T} \tag{22}$$

And the credit spread CS(T) is defined

$$CS(T) = Y^T - r \tag{23}$$

3. Data and Methodology

3.1. Sample selection

Our sample consists of all US firms with 5 year CDSs between 2003 and 2010 and all required accounting data available. This yields a total of 320 firms from a wide range of sectors. The sample period is restricted by the credit default swap spread quotes that are available on Datastream only after 2002. Although Merton's (1974) assumptions refer to a simple capital structure, we have not restricted our sample to such firms because a firm's obligations to service the debt are usually bound by contractual agreements whereby a default in one obligation would immediately trigger default in all debt obligations.[6] However, the impact of a complex capital structure requires a more careful analysis when it comes to determining each firm's recovery value. The accounting data is lagged by three months to avoid the look-ahead bias.

3.2. Model implementation

Table 1 shows the inputs that are common to all three models, and those that are specific to each.

Table 1 here

Parameters common to the three models:

- *Maturity* (T): We have used 5-year maturity CDSs, as they are most liquid.
- *Risk Free Rate* (RFR): US five-year zero rate.
- *Volatility* (σ): The asset volatility is estimated as in Bharath and Shumway (2008), i.e.:

$$\sigma = \frac{E}{E+D} \times \sigma_{Equity} + \frac{D}{E+D} \times \sigma_{Debt} \tag{24}$$

Each month from March 2003 to Setptember 2010, the equity volatility is estimated for each sample firm using daily stock returns over the past year (250 trading days) and then annualised.

Debt volatility is estimated as in Bharath and Shumway (2008):

$$\sigma_{Debt} = 0.05 + 0.25 \times \sigma_{Equity}. \tag{25}$$

- *Leverage:* Leverage is defined as the ratio of the book value of debt to the sum of the book value of debt and the market value of equity.

$$\text{Leverage} = \frac{\text{Book Value of Total Debt}}{\text{Market Value of Equity+Book Value of Total Debt}} \tag{26}$$

- *Recovery value (= 1 - loss given default)* : Debt recovery rate is defined as the payment to a class of debtholders, measured as a fraction of the face value of claims held by that debt class. Debt recovery can vary widely depending on existing creditor classes. Absolute priority requires that more junior creditors receive financial consideration in a distress restructuring only when more senior creditors are paid in full. As in Eom,

[6] Lyden and Saraniti (2001) have tested Merton (1974) empirically with firms having complex capital structure and have used the same simplification (that all debt is retired at the maturity of the bond, with equal priority given to all creditors in the event of default).

Helwege and Huang (2004), we use a recovery value of 50% of face value, which is consistent with the average defaulted debt estimated by Moodys for US entities between 1920-1999 in Keenan, Shtogrin and Sobehart (1999).

Model Specific parameters:

- *Dividend yield* (Merton and CDG) is defined as the annual dividend per share divided by the price per share. The dividend yield of the stock price is from the Datastream database.

- *Speed of adjustment* λ and *target leverage*, υ (CDG): The firm will adjust the outstanding book value in order to meet υ with a certain rate of λ. The speed of adjustment λ measures the extent to which a firm has narrowed the debt gap between last period's leverage and this period's target leverage. We use the sector average as the target mean reverting leverage.[7]

- *Global recovery* (CG): The global recovery is the average recovery of all debt obligations, assumed to be same as the recovery value.

- *Standard deviation of the default barrier* (CG): This corresponds to the standard deviation of the default barrier, which we have assumed as 0.30.

3.3. Measures of performance

We test the performance of the models using prediction error (for bias) and mean absolute prediction error (for accuracy) against observed credit spreads.

The monthly percentage prediction error for each firm in the sample is estimated as

$$\%PE_{i,t} = \frac{OS_{i,t} - IS_{i,t}}{OS_{i,t}} \tag{27}$$

The mean %PE is then estimated for each sample firm as:

$$\overline{\%PE_i} = \sum_{t=1}^{n} \frac{\%PE_{i,t}}{n} \tag{28}$$

Similarly, the monthly percentage absolute prediction error (%APE) and mean %APE are estimated as:

[7]The mean reverting leverage is calculated as the Datastream sector average using data over the period 1970-2010. The average leverage is defined as $\frac{Total\ Debt}{Total\ Debt + Market\ Value\ of\ Equity}$

$$\%APE_{i,t} = \frac{|OS_{i,t} - IS_{i,t}|}{OS_{i,t}} \tag{29}$$

$$\overline{\%APE_i} = \sum_{t=1}^{n} \frac{\%APE_{i,t}}{n} \tag{30}$$

Where:

$IS_{i,t}$ = Model implied credit spread for firm i at the end of month t.

$OS_{i,t}$ = Observed credit spread for firm i at the end of month t.

In order to formally test whether (1) model predicted spreads can explain cross-sectional variation in observed spreads, and (2) whether the model spreads under or over predict the actual spreads, we run the following Fama and MacBeth (1973) cross-sectional regressions each month:

$$\mathrm{CDS}_{\mathrm{i,t}} = \alpha_{\mathrm{t}} + \beta_{\mathrm{IS,t}}\, \mathrm{IS}_{i,t} + \varepsilon_{\mathrm{i,t}} \tag{31}$$

Where:

$CDS_{i,t}$ = actual CDS spread for firm i at the end of month t, and

$IS_{i,t}$ = Model implied credit spread for firm i at the end of month t.

If the model predicted spreads are able to explain the cross-sectional variation in observed spreads, we expect β_{IS} to be positive and statistically significant. If the model spreads are on average lower than observed spreads, we expect α to be positive, while a negative α is expected if the models over estimate the spreads.

While the implied spreads are pure measures of credit risk, the actual spreads are likely to reflect other firm specific information as well. We test this by introducing CDS liquidity and credit ratings in the pricing equation and run the following Fama and MacBeth (1973) regressions:

$$\mathrm{CDS}_{\mathrm{i,t}} = \alpha_{\mathrm{t}} + \beta_{\mathrm{RATG,t}}\, \mathrm{RATG}_{\mathrm{i,t}} + \beta_{\mathrm{LIQ,t}}\, \mathrm{LIQ}_{\mathrm{i,t}} + \beta_{\mathrm{IS,t}}\, \mathrm{IS}_{i,t} + \varepsilon_{\mathrm{i,t}} \tag{32}$$

Where:

$RATG_{i,t}$ = Standard and Poor's credit rating for firm i at the end of month t, and

$LIQ_{i,t}$ = Liquidity of the CDS of firm i at the end of month t measured as the difference between the bid and offer spread divided by the mid quote.

Other variables are as previously defined.

Following Ogden (1987) we convert credit ratings to a numerical scale from 1 (for rating AAA) to 7 (for rating CCC). We uses Moody's and Fitch's ratings (in this order) if S&P ratings are not available.

If liquidity and credit ratings do not carry any information about CDS spreads in addition to that in the model predicted spreads, we expect β_{LIQ} and β_{RATG} to be zero.

Finally, we test whether the outputs of more complex CDG and CG models carry any incremental information to that already in the simpler Merton model. We therefore introduce the predicted spreads from all three models in equations (26) and (27) simultaneously. However, the model is likely to suffer from multicollinearity due to high correlations between the spreads produced by the three models. To circumvent this problem, we use the orthogonalization procedure of Fama and French (1993). Specifically, each month, we first estimate the following regression:

$$\mathrm{CG} = \propto_1 + \gamma \mathrm{Merton} + \varepsilon_1 \tag{33}$$

We then create the variable CG(0) as:

$$\mathrm{CG}(0) = \propto_1 + \varepsilon_1 \tag{34}$$

Each month, we also estimate the following regression:

$$\mathrm{CDG} = \propto_2 + \omega_1 \mathrm{Merton} + \omega_2 \mathrm{CG}(0) + \varepsilon_2 \tag{35}$$

And then create the variable CDG(0) as

$$\mathrm{CDG}(0) = \propto_2 + \varepsilon_2 \tag{36}$$

The procedure gives CG(0) as a variable that carries information orthogonal to that in Merton, and CDG(0) as a variable that carries information orthogonal to both, Merton and CG.[8]

[8] This procedure is equivalent to the Gram-Schimdt Algorithm for constructing an orthogonal basis.

4. Results

4.1. Bias and predictive accuracy

Table 2 presents the performance analysis based on the mean prediction and absolute prediction errors. The % mean prediction errors in column 2 are positive (13.8%, 62.2%, and 36.0% for M, CDG, and CG respectively) and statistically highly significant (t=3.0, 21.2 and 7.6 respectively). Column 3 shows that the median prediction errors are also positive and statistically highly significant for all three models. The evidence is consistent with existing literature in suggesting the structural models underestimate the observed spreads, though the Merton model estimates are the least biased of the three.

Column 4 presents the % mean absolute errors and shows that the Merton model estimates are less accurate than those of other two models (114.2% for M as compared to 95.1% and 106.6% for CDG and CG respectively).[9] The large magnitude of absolute prediction errors shows that the three structural models produce downward biased estimates that can be quite inaccurate.

Table 2 here

4.2 Test of predictive ability through cross-section regressions

In order to test the ability of the structural models in explaining the cross-sectional variation in observed spreads, we run the regressions in equations (31) and (32) every month from January 2003 to September 2010. Model i in table 3 shows that the Merton model predicted spreads are strong predictors of observed spreads (t = 19.5). Similar conclusions can be drawn from models (ii) and (iii) with respect to implied spreads from CDG and CG models (t = 19.7 and 24.1 respectively). Further, model (iv) shows that while the CG models does not carry information incremental to that in M model, the CDG model

[9] The median % absolute errors are much closer with 99.2% for M as compared to 92.3% for CDG and 95.5% for CG.

does. The three models taken together are able to explain an impressive 45% of the variation in observed spreads.

However, the intercepts in all four models in table 3 are positive and statistically highly significant (smallest t value is 12.0), confirming the evidence in table 2 as well as in existing literature, structural models underestimate the observed credit spreads.[10]

Table 3 here

Delianedis and Geske (2001) and Longstaff, Mithal and Neis (2005) find that the default risk explains a much larger proportion of observed credit spreads for firms with lower credit ratings. Since we have a small number of observations in some ratings categories, we aggregate the ratings in to three categories, A or better, BBB, and BB or worse.[11] Table 4 panel A shows that while the three structural models tested here significantly underpredict the credit spreads in all rating categories, consistent with existing literature, the prediction errors are smaller for lower rating. Further, panel B shows that the absolute prediction errors are also lower for lower rating demonstrating increased prediction accuracy for such firms.

Table 4 here

The evidence presented so far demonstrates that actual credit spreads are, on average, higher than those predicted by the structural models. One potential explanation for this discrepancy is that while the implied spreads are pure measures of credit risk, the actual spreads may be influenced by other factors such as liquidity. Table 5 presents the results of Fama and MacBeth (1973) regressions when we introduce credit ratings and CDS liquidity in the model. Contrary to the findings of Ericsson et al. (2007), the evidence shows that the actual spreads contain both, additional credit risk information not captured by structural models as well as liquidity related information. While the structural model implied spreads are statistically highly significant in all four models in table 5, the coefficients on both, credit ratings (lowest t value = 9.5) as well as liquidity (lowest t value = 15.7) are positive and highly significant. The explanatory power of the models with

[10] Untabulated results show that both CG and CDG carry incremental information using implied volatility or 10 year CDS. The intercepts however, are always positive and highly significant.
[11] We have 2 observations with AAA, 14 with AA, and 6 with CCC rating.

credit rating and liquidity variables in addition to the structural model output jumps to a remarkable 76% or more. Interestingly, the intercepts are now negative and statistically highly significant in all four models (lowest t-statistic is 9.9), demonstrating that the apparent underprediction of actual spreads observed in the existing literature as well as table 3 above, is actually due to structural models missing some credit risk relevant information as well as the influence of liquidity. In fact, when liquidity and credit ratings are taken in account, the observed spreads are lower than those implied by the structural models.

Table 5 here

5. Robustness checks

5.1. Sub-period analysis

The credit crisis beginning with the collapse of Bear Sterns in March 2008 has had profound impact on the global credit markets. In this sub-section, we repeat our analyses separately for the period before and after the crisis. Results not tabulated here to save space show that before considering the impact of liquidity and credit ratings (as in table 3), the structural model spreads are lower than observed spreads in both periods. Including credit ratings and CDS liquidity measures in the regressions (as in table 5) shows, as expected, a much higher impact of liquidity on observed spreads in the post Bear Sterns period. Further, while the observed CDS spreads are, as for the full period, lower than implied spreads for both sub-periods once other firm specific variables are controlled for, the intercepts in the second sub-period are much larger in absolute terms than in the first period suggesting an even greater difference between observed and implied spreads in this period.

5.2. Alternative specifications of asset volatility

Asset volatility is a key driver of the spreads generated by the structural models. However, it is unobservable and there is no consensus in the literature on how it should be estimated. The Creditgrades^{TM} technical document (2002), for instance uses historical equity volatility over 252 to 1250 days while Eom et al. (2004) use 30 to 150 days. Cao, Yu and Zhong (2011) however claim that using option implied volatility yields better structural model estimates. We therefore repeat all our analysis with 90 days, 150 days, and 1000 days historical volatility as well as option implied volatility using average of the first available at-the-money put and call options with over 1 month maturity. All our conclusions remain unchanged with alternative specifications and are not tabulated here to save space. Our results with 10-year CDSs are also qualitatively the same.

5.3. Industry analysis

Table 6 presents the sector wise median prediction errors.[12] It shows that the underprediction of credit spreads is not driven by any particular industry, while Merton model underestimates the spreads for 25 of the 30 sectors, CDG does so for all 30 and CG for 29.

Table 6 here

6. Conclusions

This paper tests the performance of three structural models using a sample of 320 U.S. companies in a wide spectrum of sectors between 2003 and 2010. In addition to the basic Merton (1974) model, we include the Collin Dufresne and Goldstein (2001) stationary leverage ratio model as it has been shown to be the best performing structural model (e.g. Eom et al. 2004), and the industry standard CreditGrades™ (2002) model. Unlike most of the existing literature, we use CDS spreads rather than bond spreads in our empirical tests because the former provide a much cleaner proxy for default risk. In addition, we do not restrict our sample to firms with simple capital structure and this allows our sample size to be much larger than extant studies leading to more robust conclusions.

The results show that consistent with existing literature, all three structural models produce spreads that are much lower than the observed CDS spreads, and this persists in all rating categories and industry sectors. We also show that the three models produce large prediction errors in addition to the bias. However, the news for structural models is not all bad, we find evidence that structural models can explain approximately a third of the cross sectional variation in observed CDS spreads. We also find that the three models tested here carry information incremental to each other, as well as to credit ratings and liquidity. However, contrary to the evidence of Ericsson et al. (2007), we find that credit

[12] We only include those sectors for which we have at least 5 observations.

ratings and liquidity, both have a very strong influence on cross section of CDS spreads. Together with the structural models, they can capture up to 80% of the variation. Interestingly, we find that once we account for liquidity and credit ratings, the underprediction of credit spreads is replaced by overprediction, i.e., the observed credit spreads are lower than implied spreads once other firm specific factors are controlled for.

Since liquidity increases the explanatory power of implied spreads, we conclude that the observed credit spreads are not driven solely by firm default risk. Whether liquidity risk should be modelled depends on whether the ultimate goal of a structural model is to match observed spreads, or to price the firm's fundamental risk of default and convert it into a credit spread. However, the inability of our structural models to capture all the information in credit ratings indicates these models do not capture even the default risk related information completely.

Whether the difference between implied and observed credit spreads is due to other factors not-related to default risk or due to market inefficiency is a question we leave for future research.

References

BHARATH Sreedhar T, SHUMWAY Tyler. 2008: "Forecasting Default with the Merton Distance to Default Model." *The Review of Financial Studies*, v 21, n 3.

BLACK Fischer, COX, John C. 1976. "Valuing Corporate Securities – Some effects of bond indenture provisions", *The Journal of Finance.* Cambridge, vol. 31, iss. 2; (May):p. 351.

COLLIN-DUFRESNE Pierre, GOLDSTEIN Robert S. 2001."Do credit spreads reflect stationary leverage ratios?", *The Journal of Finance.* Cambridge, vol. 56, iss. 5 (Oct.):pg. 1929, 30 pgs.

EOM Young Ho , HUANG Jing-Zhi, HELWEGE Jean. 2004. "Structural Models of Corporate Bond Pricing: An Empirical Analysis", *The Review of Financial Studies*, vol.17, no.2, pp499 - 142 544.

ERICSSON Jan, RENEBY Joel, WANG Hao. 2007. "Can Structural Models Price Default Risk? Evidence from Bond and Credit Derivative Markets", *Working Paper, McGill University.*

FAMA Eugene, FRENCH Kenneth. 1993. "Common risk factors in the returns on stock and bonds", *Journal of Financial Economics* 33, 3-56 North-Holland.

FAMA Eugene F., MACBETH James D. 1973. "Risk, Return, and Equilibrium: Empirical Tests", *Journal of Political Economy,* vol. 81, no.3, 607-636.

FINKELSTEIN V, PAN George, LARDY J. P., TA Thomas, and TIERNEY John, FINGER Christopher, 2002. "Creditgrades™ Technical Document", *Riskmetrics Group.*

DELIANEDIS Gordon, GESKE Robert. 2001. "The Components of Corporate Credit Spreads: Default, Recovery, Taxes, Jumps, Liquidity, and Market Factors", *UCLA Anderson Working Paper* no. 22-01, (December).

GESKE Robert. 1979. "The Valuation of Compound Options", *Journal of Financial Economics,* Amsterdam, vol. 7, iss. 1(March):p. 63.

HUANG J., ZHOU H.: Specification Analysis of Structural Credit Risk Models. Available at http://ssrn.com/abstract=1105640.

JONES E.P, MASON.S., ROSENFELD E. 1984. "Contingent Claim Analysis of Corporate Capital Structure: An Empirical Investigation", *Journal of Finance*, 39, pp.611-627.

KEENAN, S., SHTOGRIN I, and SOBEHART. 1999. "Historical default rates of corporate bond issuers, 1920-1998," Special Comment, Moody's Investors Service.

LELAND, Hayne E. 1994. "Corporate debt value, bond covenants, and optimal capital structure", *The Journal of Finance.* Cambridge, vol. 49, iss. 4 (Sep): pg. 1213, 40pgs

LELAND, H. E.,TOFT K. B. 1996. "Optimal Capital Structure, Endogenous Bankrupcy, and the Term Structure of Credit Spreads", *Journal of Finance*, 51, pp. 987-1019.

LYDEN Scott, SARANITI David. 2001. An Empirical Examination of the Classical Theory of Corporate Security Valuation (May). Working Paper Available at SSRN.

LONGSTAFF Francis A, SCHWARTZ. Eduardo S. 1995. "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt", *Journal of Finance*, vol. 50, no. 3, (July): pp. 789-819.

MERTON Robert. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates", *Journal of Finance*, vol.29, issue 2, pages 449-470.

OGDEN Joseph P. 1987. "Determinants of the Ratings and Yields on Corporate Bonds: Tests of the Contingent Claims Model", *The Journal of Financial Research*, vol.X, no.4.

Table 1: Merton, CDG and CreditGrades™ Model Inputs

Model Inputs

<table>
<tr><th>Merton</th><th>CDG
Stationary leverage</th><th>Creditgrades™</th></tr>
<tr><td colspan="3">❑ Maturity (T)
❑ Risk Free Rate (RFR)
❑ Volatility (σ)</td></tr>
<tr><td colspan="2">❑ Dividend yield (DY)
❑ Bankruptcy costs</td><td rowspan="2">❑ Global recovery
❑ Recovery rate
❑ Leverage (D, debt-per-share)
❑ Standard deviation of the default barrier</td></tr>
<tr><td>❑ Inverse leverage ratio</td><td>❑ Initial leverage
❑ Mean reverting coefficient λ=0,18
❑ Recovery rate (=1 - ω)
❑ ν = 0,6</td></tr>
</table>

Table 2: Performance of the Merton, CDG and CG models

The table presents the mean % prediction error ((observed spread – predicted spread) / observed spread), the median % prediciton error, the mean absolute and median absolute prediction error. The mean is trimmed at a 1% level The spreads are predicted with monthly frequency using Merton (1974) model (Merton), Colin-Dufresne and Goldstein (2001) model (CDG), and the Creditgrades™ (2002) model and compared to the observed 5 year CDS spread. We use historical volatility for a 250 days horizon.

	Mean % prediction error	Median % prediction error	Mean abs. % prediction error	Median abs. % prediction error
Merton	13.8	27.6	114.2	99.2
	(2.96)	(6.29)		
CDG	62.2	72.3	95.1	92.3
	(21.15)	(13.59)		
CG	36.0	52.8	106.6	95.5
	(7.57)	(10.93)		

Table 3: Cross-sectional explanatory power of structural models

Each month from January 2003 to September 2010, the credit spreads for each firm with a 5 year CDS are generated using Merton (1974) model (Merton), Colin-Dufresne and Goldstein (2001) model (CDG), and the Creditgrades^{TM} (2002) model. In addition, CG(0) spreads are generated each month by orthogonalising CG spreads with respect to Merton spreads, and CDG(0) is generated by orthogonalising CDG spreads with respect to Merton and CG(0). The Fama and MacBeth (1973) regressions are then conducted each month with the observed CDS spreads as the dependent variable. Figures in brackets are the t-stats. We use historical volatility with a 250 days horizon.

$$\mathrm{CDS_{i,t}} = \alpha_\mathrm{t} + \beta_{\mathrm{IS,t}}\, \mathrm{IS}_{i,t} + \varepsilon_{\mathrm{i,t}} \tag{31}$$

	(i)	(ii)	(iii)	(iv)
Intercept	0.008 (13.55)	0.007 (11.99)	0.008 (16.48)	0.007 (16.67)
Merton	0.842 (19.50)			0.747 (12.71)
CDG		1.271 (19.66)		
CG			0.682 (24.07)	
CDG(0)				0.473 (7.57)
CG(0)				0.099 (1.23)
Adj. R^2	0.35	0.34	0.36	0.45

Table 4: Structural model performance by S&P rating

The table presents the median percentage prediction errors using equation 27, and the median absolute prediction error using equation 29. The median percentage prediction error and the median percentage absolute prediction error are presented for ratings in three categories: A or better rating, BBB and BB or worse rating. The implied spreads are predicted with monthly frequency using Merton (1974) model (M), Colin-Dufresne and Goldstein (2001) model (CDG), and the Creditgrades^{TM} (2002) model (CG) and compared to the observed CDS spread. Volatility is historical with a time horizon of 250 days. Figures in brackets are the z-statistics from Wilcoxon rank sum test for median = 0.

	Credit rating		
	A or better (n = 89)	BBB (n = 138)	BB or worse (n = 89)
A. Median % prediction error			
M	43.6 (3.34)	23.6 (3.71)	27.0 (4.18)
CDG	87.0 (6.36)	70.6 (9.49)	66.0 (7.25)
CG	75.3 (5.48)	58.0 (8.12)	33.0 (5.01)
B. Median % absolute prediction error			
M	112.3	104.5	85.0
CDG	98.5	93.4	81.1
CG	99.2	95.6	86.1

Table 5: Explaining the cross-section of CDS spreads

Each month from January 2003 to September 2010, the credit spreads for each firm with a 5 year CDS are generated using Merton (1974) model (Merton), Colin-Dufresne and Goldstein (2001) model (CDG), and the Creditgrades^{TM} (2002) model. In addition, CG(0) spreads are generated each month by orthogonalising CG spreads with respect to Merton spreads, and CDG(0) is generated by orthogonalising CDG spreads with respect to Merton and CG(0). RATG is the credit rating of each reference obligation each month, and LIQ is the bid-offer spread for each CDS. The Fama and MacBeth (1973) regressions are then conducted each month with the observed CDS spreads as the dependent variable. Figures in brackets are the t-stats. We use historical volatility with a 250 days horizon.

$$\mathrm{CDS}_{i,t} = \alpha_t + \beta_{\mathrm{RATG},t}\,\mathrm{RATG}_{i,t} + \beta_{\mathrm{LIQ},t}\,\mathrm{LIQ}_{i,t} + \beta_{\mathrm{IS},t}\,\mathrm{IS}_{i,t} + \varepsilon_{i,t} \qquad (32)$$

	(i)	(ii)	(iii)	(iv)
Intercept	-0.014 (-9.93)	-0.013 (-10.52)	-0.013 (-11.35)	-0.012 (-11.04)
RATG	0.002 (9.54)	0.002 (11.18)	0.002 (11.83)	0.002 (11.54)
LIQ	9.400 (15.99)	9.688 (16.80)	9.561 (16.35)	9.086 (15.66)
Merton	0.362 (15.01)			0.383 (14.22)
CDG		0.492 (17.09)		
CG			0.310 (19.72)	
CDG(0)				0.056 (1.57)
CG(0)				0.206 (3.59)
Adj. R^2	0.76	0.77	0.78	0.80

Table 6 Structural model performance by sector
The table presents the median percentage prediction errors using equation 27 and the z-statistics from Wilcoxon rank sum test for median = 0 for 31 sectors (sectors in the sample with five or more firms). The spreads are predicted with monthly frequency using Merton (1974) model (M), Colin-Dufresne and Goldstein (2001) model (CDG), and the Creditgrades^{TM} (2002) model (CG) and compared to the observed five year CDS spread. We use historical volatility with a 250 days horizon.

	Median % PE			z-statistics		
	Merton	CDG	CG	Merton	CDG	CG
Aerospace & Defense	79.4%	92.8%	93.5%	2.20	2.37	2.20
Automobiles & Parts	22.0%	61.0%	40.6%	1.86	2.37	1.35
Beverages	98.4%	99.5%	99.7%	1.76	2.06	2.06
Chemicals	13.0%	69.1%	50.3%	1.02	2.70	1.78
Construction & Materials	35.7%	86.9%	63.0%	1.61	2.37	2.03
Electricity	32.7%	62.7%	62.5%	1.19	3.41	3.41
Electronic & Electrical Equipment	38.3%	84.2%	58.3%	0.94	1.90	0.94
Financial Services	-35.6%	23.7%	-19.7%	-1.38	0.76	-1.33
Food & Drug Retailers	64.2%	93.7%	82.7%	1.96	2.85	1.96
Food Producers	76.6%	98.4%	91.8%	2.58	2.94	2.94
Gas, Water & Multiutilities	-15.9%	63.2%	69.5%	-0.34	2.20	1.86
General Industrials	39.1%	77.8%	73.9%	0.41	2.03	1.48
General Retailers	43.6%	82.1%	68.8%	3.72	2.92	3.85
Health Care Equipment & Services	77.3%	92.1%	83.4%	2.85	2.94	2.94
Household Goods & Construction	-16.3%	33.7%	7.5%	-0.11	3.10	1.34
Industrial Engineering	14.3%	63.3%	31.7%	-0.31	2.20	0.73
Industrial Metals & Mining	14.6%	56.7%	33.0%	-0.34	2.37	1.35
Industrial Transportation	19.6%	62.5%	49.5%	0.70	2.52	2.10
Life Insurance	-1.1%	58.4%	18.0%	-0.13	2.02	0.14
Media	31.1%	73.6%	39.0%	0.89	2.94	1.33
Mobile Telecommunications	44.9%	67.4%	50.5%	0.73	0.94	0.94
Nonlife Insurance	22.7%	85.8%	56.6%	0.31	3.06	1.14
Oil & Gas Producers	19.7%	60.9%	45.4%	2.05	3.82	3.78
Oil Equipment & Services	11.3%	54.5%	38.9%	-0.31	2.69	2.22
Personal Goods	92.7%	100.0%	98.8%	2.02	2.06	2.03
Pharmaceuticals & Biotechnology	87.5%	99.9%	93.9%	2.67	2.89	2.68
Real Estate Investment Trusts	-23.3%	47.8%	20.3%	-2.05	3.17	2.65
Software & Computer Services	46.9%	94.1%	48.9%	2.02	2.02	2.03
Support Services	59.0%	39.4%	55.4%	2.02	1.75	2.02
Technology Hardware & Equipment	53.2%	92.3%	69.9%	2.31	2.94	2.93
Travel & Leisure	29.9%	69.0%	53.7%	1.43	2.20	1.29