
Antje Berndt†  Darrell Duffie‡  Rohan Douglas§  Mark Ferguson¶

August 18, 2011

Abstract

*JEL Classifications:*

*Keywords: Default risk premia*

†Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA, 15213. Phone: 412-268-1871. Email: aberndt@andrew.cmu.edu.
‡Graduate School of Business, Stanford University, Stanford, CA, 94305. Phone: 650-723-1976. Email: duffie@stanford.edu.
§Quantifi Inc, Summit, NJ, 07901. Phone: 908-273-9455. Email: rdouglas@quantifisolutions.com.
¶Simon Graduate School of Business, University of Rochester, Rochester, NY, 14627. Phone: . Email: .
1 Introduction

This paper estimates the degree of variation over time in the price for bearing default risk, above and beyond expected default loss, for U.S. corporate debt. We exploit the close but time-varying relationship between default probabilities, as estimated by the Moody's KMV EDF measure, and default swap (CDS) market rates.

2 Data

This section discusses our data sources for default swap rates and for conditional default probabilities.

2.1 CDS data

We obtain daily CDS quotes from Markit for 1- and 5-year, quarterly premium, senior unsecured, U.S.-dollar-denominated, at-the-money default swaps, over the sample period from January 2001 to June 2010. Prompted by calls for standardization in the credit markets in light of the 2007-8 credit crisis and the failure of Lehman Brothers Inc., the “Big Bang” Protocol was implemented on April 8, 2009. The most visible change is the standardization of the coupon payments on CDS to fixed payments of either 100bps for investment-grade credits or 500bps for high-yield credits. For the vast majority of CDS, this will require an upfront payment by one counterparty to the other to customize the coupon rate to match the credit profile of the underlying bond. Other changes to make CDS uniform include standard quarterly settlement dates, standardized accrual terms, and a prohibition against restructuring the CDS during its term.

For the majority of the CDS quotes provided by Markit, the contractual definition of default only allows for bankruptcy and a material failure by the obligor to make payments on its debt. But a number of quotes prior to April 2009 also consider a restructuring of the firm’s debt that is materially adverse to the interests of creditors as a default event. The contractual
definition of default can affect the estimated risk-neutral implied default probabilities, since of course a wider definition of default implies a higher risk-neutral default probability. To keep the degree of heterogeneity in our data over the definition of default to a minimum, we only include CDS spreads quoted under No Restructuring. This has the additional advantage that for a given level of seniority—our data are based on senior unsecured debt instruments—there is less recovery-value heterogeneity if the event of default is bankruptcy or failure to pay only, for these events normally trigger cross-acceleration covenants that cause debt of equal seniority to convert to immediate obligations that are pari passu, that is, of equal legal priority.

We collect CDS data for all constituents of the CDX.NA.IG index (series 1 through 13), the CDX.NA.HY index (series 1 through 13), and the CDX.NA.XO index (series 5 through 11). We restrict the sample to publicly traded companies incorporated in the U.S., and only retain firms were identifiers can be matched unambiguously across Compustat, CRSP, Markit and Moody’s EDF databases. We exclude firms that have less than 180 days of 5-year CDS quotes during the sample period. This leaves us with 245 firms in nine industries.\(^1\) The range of credit qualities of the included firms may be judged from Table 1, which shows, for each industry and credit rating, the number of firms in our study of that median Moodys rating during the sample period. Table 1 indicates a concentration of A, Baa, Ba and B-rated firms.

![Table 1 about here.]

The 5-year CDS quotes are the most liquid, and are the basis for our empirical analysis.\(^2\) The rates provided by Markit are composite CDS quotes. They are computed based on

\(^1\)The eight industries are Mining, Utilities and Construction (identified as firms with NAICS codes between 210000 and 239999), Manufacturing (NAICS codes from 310000 to 339999), Trade and Transportation (NAICS codes from 420000 to 429999, 440000 to 459999, and 480000 to 499999), Information (NAICS codes from 510000 to 519999), Finance, Insurance and Real Estate (NAICS codes from 520000 to 539999), Services (NAICS codes from 540000 to 569999 and 810000 to 819999), Education and Healthcare (NAICS codes from 610000 to 629999), Entertainment and Accommodation (NAICS codes from 710000 to 729999), and Others.

\(^2\)Our main time-series analysis relies more heavily on the 5-year benchmark, allowing for noisy observation of the 1-year quotes.
quotes obtained by 2 or more anonymous sources, including investment banks and default-swap brokers. The distribution of the number of sources across all 5-year CDS quotes is shown in Figure 1. For our sample, the number of contributors ranges from 2 to 27, with the mean and median composite depth being 7.

[Figure 1 about here.]

The time series of median 5-year CDS rates for the firms in our sample is plotted in the top panel of Figure 2. Early in the sample, we see credit spreads rise sharply late in 2001, and again in the third quarter of 2002 following the Worldcom scandal. After a slow decline in CDS spreads over the next 18 months, rates stayed fairly flat from 2004 until the beginning of the financial crisis in 2007. One notable exception was the sudden increase in credit spreads from mid-March to mid-May 2005 in response to profit warnings and the eventual downgrading of Ford and General Motors to junk status. Then in mid-2007 credit spreads widened again between late June and mid-August, around the time of the collapse of two Bear Stearns hedge funds. In mid-October 2007, Citigroup and later Merrill Lynch announce record losses and write-downs on subprime-related debt and leveraged loans. Market-wide spreads continue to rise until they peak in March 2008 shortly after the Bear Stearns bailout. After a short tightening, CDS spreads begin to widen again in May and finally peak at their highest level in late 2008 following the Citigroup bailout. Median CDS rates then recovered, dipping at below 100 basis points in January 2010.

[Figure 2 about here.]

Figure 3 shows that the elevation in credit spreads during these periods occurred simultaneously across many industries, and that these episodes can be thought of as market-wide credit events.

[Figure 3 about here.]
Further confirmation for the cross-sectional comovement in credit spreads is provided in Figure 4, which shows the time series of median CDS rates, conditioned on Moody’s long-term senior-unsecured issuer rating. Holding the credit rating constant, we observed dramatic changes in credit spreads over time, with high spreads at the beginning and towards the end of our sample period and low spreads during 2004-6 period in between.

We consider the findings in Figure 4 preliminary evidence of the temporal variation in default risk premia in the following sense. In frictionless markets, the CDS rate is a close approximation of the par-coupon credit spread of the same maturity as the default swap, due to arbitrage reasoning shown by Duffie (1999).\footnote{Indeed, that par credit spreads are relatively close to CDS rates is confirmed in the empirical analysis of Blanco, Brennan, and Marsh (2005), provided one measures bond spreads relative to interest-rate swap yields, rather than treasury yields, which can be contaminated by tax exemption of coupon income, repo specials, and liquidity effects. To the extent that CDS rates differ from bond credit spreads, Blanco, Brennan, and Marsh (2005) indicate that CDS rates tend to reflect slightly fresher information.} At the same time, the average historical default frequency of firms with the same credit rating as the target firm is a common industry measure of default likelihood. Our results in Figure 4 thus speak to the relationship between default probabilities and corporate credit spreads, and provide empirical evidence of substantial temporal variation in the latter while holding the rating-based default probabilities constant.

While average default rates by credit rating are often used in practice, the ratings agencies, however, do not claim that their ratings are intended to be a measure of default probability, and they acknowledge a tendency to adjust ratings only gradually to new information, a tendency strongly apparent in the empirical analysis of Behar and Nagpal (1999), Lando and Skdeberg (2002), Kavvathas (2001), and Nickell, Perraudin, and Varotto (2000), among others.\footnote{Average historical default probabilities by credit rating are used, for example, in implementations of the CreditMetrics approach (www.creditmetrics.com). They are convenient given the usual practice by financial-services firms of tracking credit quality by internal credit ratings based on the approach of the major recognized rating agencies such as Moodys and Standard and Poors.} This tendency to adjust ratings gradually is illustrated in Moodys (2004), which
shows dramatic variation in default rates by rating depending on whether the prior rating was higher or lower.

Bohn, Arora, and Korablev (2005) report that ratings-based default prediction has an out-of-sample accuracy ratio for 2000-2004 of 0.72. The authors also show that rating-based predictions are clearly outperformed by expected default frequency (EDF) estimates as provided by Moody’s KMV, which have an accuracy ratio of 0.84 for the same time span.

2.2 EDF data

Moodys KMV provides its customers with, among other data, current firm-by-firm estimates of conditional probabilities of default over time horizons that include the benchmark horizons of 1 and 5 years. For a given firm and time horizon, this EDF estimate of default probability is fit non-parametrically from the historical default frequency of other firms that had the same estimated distance to default as the target firm.\(^5\) The distance to default of a given firm is a leverage measure adjusted for current market asset volatility. Roughly speaking, distance to default is the number of standard deviations of annual asset growth by which the firms expected assets at a given maturity exceed a measure of book liabilities. The liability measure is, in the current implementation of the EDF model, the firms short-term book liabilities plus one half of its long-term book liabilities. Estimates of current assets and the current standard deviation of asset growth (volatility) are calibrated from historical observations of the firms equity-market capitalization and of the liability measure. The calibration, explained for example in Vassalou and Xing (2004), is based on the model of Black and Scholes (1973) and Merton (1974), by which the price of a firms equity may be viewed as the price of an option on assets struck at the level of liabilities. Crosbie and

---

\(^5\)Moody’s definition of default used for purposes of the Moodys KMV EDF estimator of default probability includes restructuring of the firm’s debt that is materially adverse to the interests of creditors in addition to bankruptcy and failure to pay. While the probability of default widens as more events are covered, Berndt, Jarrow and Kang (2007) show that the likelihood of a restructuring event is substantially smaller than that of a bankruptcy or missed payment. The EDF measure therefore assigns a slightly higher probability than that for the events covered by the CDS contracts in our sample. While it may bias our results towards lower default risk premia estimates, we believe that the effect is minimal. Nevertheless, we provide a number of robustness checks in Section xxx
Bohn (2002) and Kealhofer (2003) provide more details on the KMV model and the fitting procedures for distance to default and EDF. Bharath and Shumway (2001) show that the fitting procedure is relatively robust.

While one could criticize the EDF measure as an estimator of the true conditional default probability, it has some important merits for business practice and for our study, relative to other available approaches to estimating conditional default probabilities. First, the Moodys KMV EDF is readily available for essentially all public U.S. companies. The EDF is fitted non-parametrically to the distance to default, and is therefore not especially sensitive, at least on average, to model mis-specification. While the measured distance to default is itself based on a theoretical option-pricing model, the function that maps DD to EDF is consistently estimated in a stationary setting, even if the underlying theoretical relationship between DD and default probability does not apply. That is, conditional on only the distance to default, the measured EDF is equal to the true DD-conditional default probability as the number of observations goes to infinity, under typical mixing and other technical conditions for non-parametric qualitative-response estimation.

The Moodys KMV EDF measure is also extensively used in the financial services industry. For example, from information provided to us by Moodys KMV, 40 of the worlds 50 largest financial institutions are subscribers. Indeed, Moodys KMV is the most widely used name-specific major source of conditional default probability estimates of which we are aware, covering over 26,000 publicly traded firms.

The middle panel of Figure 2 displays the time series of median 1-year and 5-year EDFs. While EDFs do increase market wide both in 2001 and 2002, and then again during the recent financial crisis, it is interesting to point out they decreased steadily in between, from 1-year EDFs at about 50 basis points in 2003 to less than 10 basis points in mid-2007.

Our basic analysis in Section 3 directly relates daily observations of 5-year CDS rates to the associated daily 5-year EDF observations. For our time-series model of default intensities, however, we turn in Section xxx to 1-year EDFs. By using 1-year EDFs rather than 5-year
EDFs, our intensity estimates are less sensitive to model mis-specification, as the 1-year EDF is theoretically much closer to the intensity than is the 5-year EDF. As a robustness check, we have also fit our time-series model of default risk premia to 5-year EDF data; the results are similar in major respects to those reported here.

2.3 Default hazard rates

Duffie, Saita, and Wang (2007) describe a more elaborate default prediction model than that employed by Moody’s KMV, using distance to default as well as other firm-specific and macro-economic explanatory variables, that achieves an accuracy ratio that is slightly higher than that of the EDF during this period. Lando and Nielsen (2009) extend the DSW model by including additional conditioning variables. Using the parameter estimates from both papers, we compute default hazard rates for each firm in our sample. The medians are shown in the bottom panel of Figure 2. They exhibit a similar temporal pattern as that observed for EDFs, but are generally of a lower magnitude.

3 Descriptive Analysis of CDS and EDF Data

In order to obtain a simple and relatively robust measure of the sensitivity of credit spreads (CDS rates) to default probabilities, we undertook a panel-regression analysis of all 426,825 paired CDS and EDF observations. A simple preliminary ordinary-least-squares (OLS) linear model of the relationship between a firm’s 5-year CDS rate and its annualized 5-year EDF, measured in basis points on the same day, reveals that the CDS rate increases on average by roughly 14 basis points for each 10 basis point increase in the 5-year EDF. If one were to take the risk-neutral expected loss given default to be, say, 60% and the annual conditional default probabilities to be constant over time, this would imply an average ratio of risk-neutral to actual annual default probabilities of approximately \((14/0.6)/10\), or 2.3, roughly consistent with the results of Driessen (2005). The associated coefficient of determination,
$R^2$, is 45.8%.

Linearity of the CDS-EDF relationship, however, is placed in doubt by a sizable intercept estimate of roughly 74 basis points, more than 90 times its standard error. Absent an unexpectedly large liquidity impact on CDS rates, the fitted default swap rate should be closer to zero at low levels of EDF. Scatter plots of the CDS-EDF relationship also reveal a pronounced concavity at low levels of EDF. That is, the sensitivity of credit spreads to a firm's default probability seems to decline as default probabilities increase. There is also apparent heteroskedasticity, with dramatically greater variance for higher EDFs. The slope of the fit is thus heavily influenced by the CDS-to-EDF relationship for lower-quality firms.

In order to mitigate the effects of non-linearity and heteroskedasticity, we considered the log-log specification

$$\log Y_i = \alpha + \beta \log X_i + z_i,$$

where $(Y_i, X_i)$ is the $i$-th observed matched pair of 5-year CDS rate and 5-year EDF for the same firm on a given date, for coefficients $\alpha$ and $\beta$, and a residual $z_i$. The fit, which has an $R^2$ of 0.62, is illustrated in Figure ??, showing much less heteroskedasticity.\(^6\) One might have considered a model in which the CDS rate is fit to both 5-year and 1-year EDF observations, given the potential for additional influences of near-term default risk on CDS rates. We have found, however, that the 1-year and 5-year EDFs are extremely highly correlated. As might be expected, adding 1-year EDFs to the regression has no major impact on the quality of fitted CDS rates, and involves substantial noise in the slope coefficients.

[Figure 5 about here.]

We also control for changes in the CDS-to-EDF relationship across time and across sectors. The manufacturing sector for June 2010 is the reference sector and month. With

\(^6\)The granularity at EDFs of 1 basis point is associated with the the floor imposed by Moody’s KMV on 5-year EDFs.
an R2 of 71.6%, the fitted model for the oil-and-gas sector may be summarized as

\[
\log Y_i = 2.0821 + 0.6745 \log X_i + \sum_{j,k} \hat{\beta}_{j,k} D_{j,k}(i) + \varepsilon_i, \tag{2}
\]

where \( \hat{\beta}_{j,k} \) denotes the estimate for the dummy multiplier for month \( j \) and sector \( k \). \( D_{j,k}(i) \) is 1 if observation \( i \) is from month \( j \) and sector \( k \) and zero otherwise, and \( \varepsilon_i \) denotes the residual.

The standard-error estimates reported here are robust to heteroskedasticity and correlation of disturbances, using the usual generalized-least-squares estimator for the covariance matrix of regressor coefficients for panel-data regressions, found, for example, in Woolridge (2002), Section 7.8.4.

The standard error for (3) of approximately 0.63, and an assumption of normally distributed disturbances, imply a one-standard-deviation confidence band for a given CDS rate of between 53% and 188% of the fitted rate. While the CDS data are noisy in this sense, the relationship between CDS and EDF is highly significant, and variation in EDF on its own explains a large fraction (an \( R^2 \) of about 62%, before controlling for time and sector effects) of variation in CDS rates. Figure 6 shows, for each sector \( k \), variation over month \( j \) of \( e^{\hat{\beta}_{j,k}} \), an estimate of the proportional variation over time of risk premia. That is, \( e^{\hat{\beta}_{j,k}} \) is the ratio of the fitted default swap rate for a firm in sector \( k \) at month \( j \), to that of a manufacturing firm with the same default probability in June 2010.

[Figure 6 about here.]
Table 1: **Descriptive Statistics** The table reports the distribution of median Moody’s issuer ratings within each sector. The sample period is January 2001 to June 2010. For a given firm, the median rating is computed over those days in the sample period for which both a letter rating and credit pricing data were available.

<table>
<thead>
<tr>
<th>Sector</th>
<th>Aaa</th>
<th>Aa</th>
<th>A</th>
<th>Baa</th>
<th>Ba</th>
<th>B</th>
<th>Caa</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education &amp; Healthcare</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Entertainment &amp; Accommodations</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Finance, Insurance &amp; Real Estate</td>
<td>3</td>
<td>8</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Information</td>
<td>6</td>
<td>9</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>28</td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1</td>
<td>20</td>
<td>34</td>
<td>21</td>
<td>14</td>
<td>3</td>
<td>1</td>
<td>90</td>
</tr>
<tr>
<td>Mining, Utilities &amp; Construction</td>
<td>15</td>
<td>9</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34</td>
</tr>
<tr>
<td>Services</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>Trade &amp; Transportation</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>21</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>42</td>
</tr>
<tr>
<td>Others</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>All</td>
<td>1</td>
<td>6</td>
<td>41</td>
<td>92</td>
<td>61</td>
<td>37</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>245</td>
</tr>
</tbody>
</table>
Figure 1: Distribution of composite depth for 5-year CDS quotes.
Figure 2: The top panel shows median 1- and 5-year CDS rates for the firms in our sample, the middle panel shows median 1- and 5-year EDF rates, and the bottom panel show median hazard rates beaded on Duffie, Saita and Wang (2007) and Lando and Nielsen (2009). To compute median rates on a given date, we require data for 30 or more firms on that date.
Figure 3: Median 5-year CDS rates for each industry.
Figure 4: Median 5-year CDS rates by credit rating.
Figure 5: Scatter plot of EDF and CDS observations in basis points, logarithmic, and OLS fitted relationship. Source: Markit (CDS) and Moody’s KMV (EDF).
Figure 6: The multipliers for estimated 5-year CDS rates, over time and sector, at a fixed 5-year EDF. These are the exponentials of the dummy coefficients in the log-CDS-to-log-EDF model (3).