Asset Volatility

Maria Correia
London Business School
mcorreia@london.edu

Johnny Kang
AQR Capital Management LLC
johnny.kang@aqr.com

Scott Richardson
AQR Capital Management LLC
London Business School
scott.richardson@aqr.com

June 30, 2014

Abstract

Asset volatility is a key variable in understanding credit risk. We evaluate alternative measures of asset volatility using information from both market (i.e., historical equity and credit market returns and equity option markets) and accounting (i.e., financial statements) sources. For a large sample of U.S. firms, we find that combining information about asset volatility from market and accounting sources improves the explanatory power of corporate bankruptcy models and cross-sectional variation in credit spreads. Market based (accounting) measures of asset volatility appear to reflect systematic (idiosyncratic) sources of volatility, and combining both sources of information generates a superior measure of total asset volatility.

JEL classification: G12; G14; M41
Key words: credit spreads; volatility; bankruptcy; default.

We are grateful to Arne Staal and Philippe Vannerem at Barclays Capital for providing us access to the Barclays Capital bond dataset. We also thank Korcan Ak, Cliff Asness, Max Bruche, John Campbell, John Hand, Jens Hilscher, Ronen Israel, Anya Kleymenova, Sonia Konstantinidi, David Lando, Lasse Pedersen, Caroline Pflueger, Tjomme Rusticus, Frank Zhang and seminar participants at American Accounting Association Annual Meeting, Cass Business School, Copenhagen Business School, European Accounting Association Annual Meeting, London Business School, LUBRAFIN and University of Porto for helpful discussion and comments. AQR Capital Management LLC invests in, among other strategies, securities studied in this paper. The views and opinions expressed herein are those of the authors and do not necessarily reflect the views of AQR Capital Management, LLC ("AQR") its affiliates, or its employees. This information does not constitute an offer or solicitation of an offer, or any advice or recommendation, by AQR, to purchase any securities or other financial instruments, and may not be construed as such.

1. Introduction

Our objective is to compare and contrast alternative measures of asset volatility in their ability to predict bankruptcy. A seminal paper by Merton (1974) developed structural models as a benchmark to describe credit risk. In these models asset volatility is arguably the most important primitive variable for determining distance to default. There is a rich literature examining how accounting data can be used to help forecast corporate bankruptcy and default (see e.g., Beaver, 1966; Altman, 1968; Ohlson, 1980; Beaver, McNichols and Rhie, 2005; Bharath and Shumway, 2008; Campbell, Hilscher and Szilagyi, 2008; and Correia, Richardson and Tuna, 2012). While many of these studies use a mix of accounting and market based variables to predict bankruptcy, a common theme in this past research is the primary use of market based measures of asset volatility. Our focus is on whether information from the accounting system could be additive to market based measures of asset volatility.

To appreciate the potential importance of accounting based measures of asset volatility, we must first note the role of asset volatility as a construct in structural models of credit risk (e.g., Merton, 1974). While there are many variants of structural models, a common theme is that a firm will 'default' if its asset value is below a default threshold at some future point. Thus, structural models provide a framework to quantify the probability that a firm will have an insufficient asset value to satisfy its current debt commitments. A firm's 'closeness' to the default threshold is a function of both (i) the expected difference between asset values and current debt commitments, and (ii) asset volatility. For a given capital structure today, higher expected asset volatility implies a greater probability that future asset values will be insufficient to cover debt commitments (i.e., a greater chance of default).

How might accounting based measures of asset volatility be additive to market based measures of asset volatility? General purpose financial reports are prepared under GAAP with the primary aim of providing information to investors (both current and prospective) to make informed investment decisions regarding their scarce economic capital. Investors are concerned not only about expected returns, but also the risk that those returns will not be achieved. While the financial reporting system is designed to record transactions about expected returns and risk, the manner in which transactions are systematically recorded means that a lot of the information recorded in financial statements is directly attributable to the specific operating and investing decisions made by that firm. Indeed, Vuolteenaho (2002) shows that decomposing firm-level return volatility into cash flow news and discount rate news components also separates return volatility into firm-specific and systematic components, respectively. We therefore expect accounting (market) based measures of asset volatility to capture relatively more idiosyncratic (systematic) sources of asset volatility. Combining both measures of asset volatility is expected to better explain bankruptcy and credit risk more generally.

In credit markets it is <u>total</u> asset volatility that matters. Irrespective of whether the source of volatility is systematic or idiosyncratic, it is critical to measure both to be able to determine whether future asset values will fall below the default threshold. Limiting measures of asset volatility to only systematic sources will generate inferior forecasts of default probability. Thus, in understanding credit risk, to the extent that market based and accounting based measures of asset volatility reflect different components of asset volatility, they are likely to both be relevant measures of asset volatility.

We source our market based measures of asset volatility from traded security prices in secondary markets. We derive several measures of historical asset volatility ranging from simplistic deleveraging of historical equity volatility to a complete measure that uses historical return volatilities and historical return correlations (see e.g., Schaefer and Strebulaev, 2008). We also combine forward looking market information using the implied volatility from at-the-money put and call options. Our accounting based measures of asset volatility are obtained from the primary financial statements and are designed to capture fundamental volatility in unlevered profitability.

We find that combining information about asset volatility from market based and accounting based information improves estimates of corporate bankruptcy. Using a large sample of firms with liquid corporate bond data, we find that a one standard deviation change in our market (accounting) based component measures of asset volatility translates to an increase of 4.2 (4.4) percent in the conditional probability of bankruptcy. Additional analysis based on the Classification and Regression Trees (CART) methodology, which allows for non-linear and interactive associations between probability of default and different explanatory variables, also shows the joint importance of accounting and market based measures of asset volatility.

While bankruptcies are relatively rare, extreme events, from an investment perspective, we show that the benefit of asset volatility in predicting bankruptcies extends to explaining the cross-section of credit spreads as well. Specifically, we find that combining market based and accounting based estimates of asset volatility improves explanatory power of cross-sectional credit spread regression models. In constrained regression analysis where we combine market and accounting based measures of asset volatility into theoretically justified implied spreads, we find that both component measures of asset volatility are relevant. Specifically, implied spreads based on market (accounting) measures of asset volatility explain 30.8 (21.2) percent of the observed variation in credit spreads.[1] We also find evidence that the relative importance of accounting based measures of asset volatility is greater for high yield corporate bonds relative to investment grade bonds.

Prior studies have documented an association between idiosyncratic equity (Campbell and Taksler, 2003) and bond (Gemmil and Keswani, 2011) volatility and credit spreads at an aggregate and issuer level. Taking these findings into account and to help better understand the relative importance of market based and accounting based component measures of asset volatility, we explore the mapping of these respective measures to systematic and idiosyncratic volatility. We find that average within industry pairwise correlations of market asset returns (a leverage-weighted average of equity and credit market returns) are significantly larger than within industry pairwise correlations of changes in seasonally adjusted accounting rates of return. This suggests that market based measures of asset volatility are more likely to reflect systematic sources of volatility. Furthermore, the first principal component of market asset returns explains 29.5% of the variation in these returns, while the first principal component of fundamental returns only explains 7.26% of the variation in fundamental returns. Together this evidence suggests that combining market based and accounting based measures of asset volatility yields a superior measure of asset

[1] These statistics are based on the R^2 Shapley decomposition (un-tabulated).

volatility due to a combination of systematic and idiosyncratic information. As discussed earlier, in the context of credit derivatives total asset volatility is the relevant measure, not just systematic volatility.

The rest of the paper is structured as follows. Section 2 describes our sample selection and research design. Section 3 presents our empirical analysis and robustness tests, and section 4 concludes.

2. Sample and research design

2.1 Secondary credit market data

Our analysis is based on a comprehensive panel of U.S. corporate bond data, which includes all the constituents of (i) Barclays U.S. Corporate Investment Grade Index, and (ii) Barclays U.S. High Yield Index. The data includes monthly returns and bond characteristics from September 1988 to February 2013. We exclude financial firms with SIC codes between 6000 and 6999.

2.2 Representative bond

Given that corporate issuers often issue multiple bonds and that our analysis is directed at measuring asset volatility of the issuer, we need to select a representative bond for each issuer. To do this, we follow the criteria in Haesen, Houweling and VanZundert (2012). We repeat this exercise every month for our sample period. The criteria used for identifying the representative bond are selected so as to create a sample of liquid and cross-sectionally comparable bonds. Specifically, we select representative bonds on the basis of (i) seniority, (ii) maturity, (iii) age, and (iv) size.

First, we filter bonds on the basis of seniority. Because most companies issue the majority of their bonds as senior debt, we select only bonds corresponding to the largest rating of the issuer. To do this we first compute the amount of bonds outstanding for each rating category for a given issuer. We then keep only those bonds that belong to the rating category which contains the largest fraction of debt outstanding. This category of bonds tends to have the same rating as the issuer. Second, we then filter bonds on the basis of maturity. If the issuer has bonds with time to maturity between 5 and 15 years, we remove all other bonds for that issuer from the sample. If not, we keep all bonds in the sample. Third, we then filter bonds on the basis of time since issuance. If the issuer has any bonds that are at most two years old, we remove all other bonds for that issuer. If not, we keep all bonds from that issuer in the sample. Finally, we filter on the basis of size. Of the remaining bonds, we pick the one with the largest amount outstanding.[2]

Our resulting sample includes 121,270 unique bond-month observations, corresponding to 5,367 bonds issued by 1,504 unique firms. Table 1, Panel A shows the industry composition of the sample, using Barclays Capital's industry definitions. Approximately 35% of the sample firms are consumer products firms. Capital Goods firms and Basic Industry make up another 20% of the sample.

---

[2] For example Basic Energy Services has two bonds in the Barclays Capital bond sample with return information for October 2009, one with rating BA3, another with rating CAA1. We first compute the fraction of debt outstanding for each rating. In this case, one half of the debt is rated BA3, and the other half CAA1, as the bonds have the same amount outstanding of 225,000. Therefore both bonds are kept in the sample after the first step. The second selection step is based on years to maturity. The first bond has 4.75 years to maturity and the second bond 6.46. We drop the first bond as time to maturity is lower than 5, and therefore the second bond is selected as the representative bond. Viacom Inc. has five bonds in the sample in December 2012, all with the same rating of BAA1. Two of these bonds have time to maturity between 5 and 15 years, therefore, we remove the remaining three bonds from the sample. Both bonds were issued at the same time. They are both 1.36 years old. Therefore, we select the representative bond based on amount outstanding.

Sample bonds have an average option adjusted spread (OAS) of 3.31% over the sample period, and an average option adjusted duration of 5.16 years (Table 1, Panel B).

2.3 Measures of asset volatility

2.3.1 Historical market data

We calculate historical equity volatility using the annualized standard deviation of CRSP realized daily stock returns over the past 252 days, σ_E. We use market leverage to de-lever historical equity volatility and obtain our first measure of asset volatility:

$$\sigma_A^{NAIVE} = \sigma_E \omega \tag{1}$$

with $\omega = \frac{E}{E+STD+LTD}$, where E is the market value of the firm's equity, STD is the book value of short term debt (Compustat mnemonic 'DLCQ'), and LTD is the book value of long term debt (Compustat mnemonic 'DLTTQ').

Our second estimate of historical asset volatility, σ_A^{ω}, combines historical credit and equity market data:

$$\sigma_A^{\omega} = \sqrt{\omega^2 \sigma_E^2 + (1-\omega)^2 \sigma_D^2 + 2\omega(1-\omega)\rho_{D,E}\sigma_E\sigma_D} \tag{2}$$

where ω is defined as in equation (1) as the fraction of asset value attributable to equity, σ_D is the annualized standard deviation of total monthly bond returns and $\rho_{D,E}$ is an estimate of the historical correlation between equity and bond returns. We compute the correlation between equity and bond returns for each bond in the representative sample over a 12-month period. Note that while our selection of a representative bond can change each month for a given issuer, our correlation and volatility measures hold a given bond fixed when looking back in time.

To mitigate noise in our estimate of historical correlations, we shrink our estimate of correlation to the average correlation for a given level of credit risk (see e.g., Lok and Richardson, 2011). Specifically, we compute $\rho_{D,E}$ for each issuer as the average correlation for all firms in the same decile of option adjusted credit spread.

In untabulated robustness tests, we also examine a Merton distance to default model estimate of asset volatility, following Bharath and Shumway (2008). This volatility is computed by simultaneously solving the Black-Scholes-Merton system of equations, using a numerical algorithm with historical equity volatility as a starting point.

Table 1, Panel B presents descriptive statistics for the variables used to de-lever volatility. Sample firms have an average market leverage of approximately 36% (1-0.6348) and exhibit an average correlation between equity and debt returns $\rho_{D,E}$ of 0.2195. Appendix I defines these variables, as well as other variables used in the paper, in more detail.

2.3.2 Forward looking market data

We obtain implied Black-Scholes volatility estimates for at-the-money 91-day options from the OptionMetrics Ivy DB standardized database.[3] We average the implied volatility for a 91-day put and call option. Based on this implied equity volatility, σ_I, we compute two asset volatility estimates, σ_{AI}^{NAIVE} and σ_{AI}^{ω}, using the approaches in (1) and (2), respectively. Option implied volatility has been shown to have incremental power with respect to historical volatility in explaining time-series and cross-sectional variation in credit spreads (Cremers, Driessen, Maenhout and Weinbaum, 2008; Cao, Yu and Zhong, 2010).

[3] The standardized implied volatilities are calculated by OptionMetrics using linear interpolation from their Volatility Surface file.

2.3.3 Fundamental data

We use two approaches to compute measures of fundamental volatility. Both approaches are designed to capture the volatility of unlevered profitability. Following Penman (2014), we use return on net operating assets ($RNOA$) as the measure of unlevered (or enterprise) profitability. First, we construct two simple measures based on historical volatility of seasonally adjusted $RNOA$, and on the volatility of $RNOA$ averaged across all four fiscal quarters. Second, we use the quantile regression approach described in Konstantinidi and Pope (2012) and Chang, Monahan and Ouazad (2013).

2.3.3.1 Naïve approach

For each quarter we compute $RNOA$ as operating income ('OIADPQ') to average NOA during the quarter. NOA is defined as the sum of common equity, preferred stock, long term debt, debt in current liabilities and minority interests minus cash and short term investments (Compustat mnemonics 'CEQQ'+'PSTKQ'+'DLTTQ'+DLCQ'+'MIBQ'-'CHEQ'). We estimate the volatility of $RNOA$, σ_F^{NAIVE}, as the standard deviation of seasonally adjusted $RNOA$ over the previous 5 years (20 quarters), requiring at least 10 available quarterly observations. Seasonally adjusted $RNOA$ for quarter k in year t is computed as:

$$RNOA_{itk}^{SA} = RNOA_{itk} - RNOA_{it-1k} \tag{3}$$

We then compute the standard deviation of seasonally adjusted $RNOA$ over the previous 5 years, requiring a minimum of 10 quarters of data.

$$\sigma_F^{NAIVE} = Std(RNOA_{itk}^{SA}) \tag{4}$$

We compute an alternative measure of $RNOA$ volatility, $\sigma_F^{NAIVE(2)}$ as:

$$\sigma_F^{NAIVE(2)} = \sum_{k=1}^{4} \frac{Std_k(RNOA_{itk})}{4} \quad (5)$$

where $Std_k(RNOA_{itk})$ is the standard deviation of $RNOA$ for quarter k calculated over the previous 20 years, requiring a minimum of 10 years of data. We annualize σ_F^{NAIVE} and $\sigma_F^{NAIVE(2)}$, by multiplying the measures by $\sqrt{4}$.[4]

2.3.3.2 Quantile regression approach

We use quantile regressions to estimate the quantiles and conditional moments of the $RNOA$ distribution. Following Konstantinidi and Pope (2012) and Chang, Monahan and Ouazad (2013), we exclude financial firms with SIC codes 6000 to 6999. We estimate coefficients for each percentile using an expanding window starting in 1963. In particular, for each year t, we estimate the following regression, using quarterly data from 1963 to t:

$$QUANT_q(RNOA_{it}|\cdot) = \beta_{0t}^q + \beta_{1t}^q RNOA_{it-1} + \beta_{2t}^q LOSS_{it-1} + \beta_{3t}^q (LOSS_{it-1} \times RNOA_{it-1}) + \beta_{4t}^q ACC_{it-1} + \beta_{5t}^q PAYER_{it-1} + \beta_{6t}^q PAYOUT_{it-1} \quad (6)$$

A full description of the explanatory variables and quantile estimation approach is provided in Appendix II.

2.3.4. Correlations across asset volatility measures

Table 1, Panel C reports descriptive statistics for the different volatility measures. We winsorize all measures of asset volatility at the 1st and 99th percentile values of their respective distributions. These measures exhibit differences in scale. In particular, volatility

[4] As an alternative naïve approach, we estimate a time-series model for $RNOA$ and calculate the time-series volatility only for the residual (the stationary component). In particular, we estimate the following regression for each firm which allows for a fixed effect across fiscal quarters: $RNOA_{itk} = \beta_0 + \beta_1 RNOA_{itk-1} + \sum_{q=2}^{4} \beta_q I_{itk}^q + \varepsilon_{it}$, where I_{itk}^q is an indicator variable equal to one if k=q. Results are very similar to those tabulated for the other naïve measures.

measures based on financial statement information, σ_F^{NAIVE}, $\sigma_F^{NAIVE(2)}$, and σ_F, are lower, on average, than asset volatility measures based on naïve or weighted deleveraging of historical equity returns or implied equity volatility. We discuss how we deal with differences in scale when using different measures of asset volatility to derive implied credit spreads in section 3.2.2.

Panel D of Table 1 reports the average monthly pairwise correlations across volatility measures. Historical equity volatility, σ_E, is highly correlated with implied volatility, σ_I, [0.8816 (0.9008) Pearson (Spearman) correlation]. The Pearson (Spearman) correlation between these equity volatility measures and debt volatility, σ_D, ranges between 0.3813 and 0.4874 (0.2651 and 0.3371), respectively. As a result, the correlations between weighted asset volatilities and the corresponding equity volatility measures are, on average, lower than 0.80.

Correlations between accounting based (σ_F, σ_F^{NAIVE}, $\sigma_F^{NAIVE(2)}$ and $P95P5$) and market based asset volatility measures ($\sigma_A^{\omega}, \sigma_A^{NAIVE}, \sigma_{AI}^{\omega}$ and σ_{AI}^{NAIVE}) are much lower. The maximum pairwise Pearson (Spearman) correlation between these two types of measures is 0.3286 (0.4339) and the minimum 0.2254 (0.2754). The two quantile-based fundamental volatility measures, σ_F and $P95P5$, exhibit a correlation close to 1, and the two naïve volatility measures, σ_F^{NAIVE}, $\sigma_F^{NAIVE(2)}$, are also highly correlated, with a Person (Spearman) correlation of 0.6356 (0.6323). However, the Person (Spearman) correlations between naïve and quantile based volatilities are much lower on average 0.4124 (0.2377). Thus, while accounting and market based measures of asset volatility are correlated they are far from perfectly correlated suggesting that they are capturing different aspects of volatility.

2.4 Bankruptcy data and distance to default

We estimate the probability of bankruptcy based on a large sample of Chapter 7 and Chapter 11 bankruptcies filed between 1980 and the end of 2012. We combine bankruptcy data from four main sources: Beaver, Correia, and McNichols (2012)[5]; the New Generation Research bankruptcy database (bankruptcydata.com); Mergent FISD; and the UCLA-Lo Pucki bankruptcy database.

We estimate probabilities of bankruptcy by using a discrete time hazard model and including three types of observations in the estimation: nonbankrupt firms, years before bankruptcy for bankrupt firms, and bankruptcy years (Shumway, 2001). Our dependent variable is equal to 1 if a firm files for bankruptcy within 1 year of the end of the month, and 0 otherwise. We keep the first bankruptcy filing and remove from the sample all months after this filing.

Following Correia, Richardson and Tuna (2012) we use quarterly financial data to compute the default barrier and update market data on a monthly basis to obtain monthly estimates of the probabilities of bankruptcy. Market variables are measured at the end of each month and accounting variables are based on the most recent quarterly information reported before the end of the month. We winsorize all independent variables at 1% and 99%. We ensure that all independent variables are observable before the declaration of bankruptcy. Furthermore, to ensure that prediction is made out of sample and to avoid a potential bias of ex post over-fitting the data, we estimate coefficients using an expanding window approach.

[5] Beaver, Correia, and McNichols (2012) combine the bankruptcy database from Beaver, McNichols, and Rhie (2005), which was derived from multiple sources including CRSP, Compustat, Bankruptcy.com, Capital Changes Reporter, and a list provided by Shumway with a list of bankruptcy firms provided by Chava and Jarrow and used in Chava and Jarrow (2004).

We convert the different scores into probabilities as follows: $Prob = e^{score}/1 + e^{score}$. All of the models are nonlinear transformations of various accounting and market data.

The primary regression model for estimating bankruptcy over the next 12 months is as follows:

$$\Pr(Y_{it+1} = 1) = f\left[ln\left(\frac{V_{it}}{X_{it}}\right), Exret_{it}, ln(E_{it}), \sigma_{k,it}\right] \quad (7)$$

The dependent variable is equal to 1 if the firm filed for bankruptcy within the following year. $ln\left(\frac{V_{it}}{X_{it}}\right)$ is a measure of dollar distance to default barrier (akin to an inverse measure of leverage). We compute V_{it} as the sum of the market value of the firm's equity and the book value of debt. We compute our default barrier, X_{it}, as the sum of short-term debt ('DLCQ') and half of long-term debt ('DLTTQ') as reported at the most recent fiscal quarter (see e.g., Bharath and Shumway, 2008). $Exret_{it}$ is the excess equity return over the value weighted market return over the previous 12 months. $ln(E_{it})$ is the logarithm of the market value of equity measured at the start of the forecasting month. $\sigma_{k,it}$ is the respective measure of asset volatility as defined in section 2.3. We estimate equation (7) using various combinations of our measures of asset volatility over different samples to assess the relative importance of market based and accounting based measures of asset volatility in the context of forecasting bankruptcy.

Our priors for equation (7) are as follows: (i) $ln\left(\frac{V_{it}}{X_{it}}\right)$ is expected to be negatively associated with bankruptcy likelihood (the further the market value of assets is from the default barrier the lower the likelihood of hitting that barrier in the next 12 months), (ii) $Exret_{it}$ is expected to be negatively associated with bankruptcy likelihood (assuming there is information content in security prices, decreases in security prices should be associated with increased bankruptcy likelihood), (iii) $ln(E_{it})$ is expected to be negatively associated with bankruptcy likelihood (this is a well-known empirical relation but the ex-ante justification is less clear; some argue that large firms offer better diversification and better realizations of asset values in the event of default), and (iv) $\sigma_{k,it}$ is expected to be positively associated with bankruptcy likelihood (the greater the volatility of the asset value the greater the chance of passing through the default barrier).

2.5 Credit Spreads

Given that a measure of asset volatility is useful in forecasting bankruptcy, and under the assumption that security prices in the secondary credit market are reasonably efficient, we also test how different combinations of measures of asset volatility are able to explain cross-sectional variation in credit spreads. We view the analysis of credit spreads as supporting evidence for assessing the information content of accounting and market based measures of asset volatility.

We do this via two approaches. First, we estimate an unconstrained cross-sectional regression where we include multiple measures of determinants of credit spreads in a linear model. Second, we estimate a constrained cross-sectional regression where we combine our various measures of asset volatility into measures of distance to default which are in turn mapped to an implied credit spread. A benefit of the constrained approach is that it combines the dollar distance to default, $ln\left(\frac{V_{it}}{X_{it}}\right)$, with measures of asset volatility, $\sigma_{k,it}$, to better identify 'closeness' to the default threshold. An unconstrained regression is unable to capture the inherent non-linear relations between leverage, asset volatility, defaults (bankruptcy) and credit spreads.

For the unconstrained approach we estimate the following regression model:

$$OAS_{it} = \alpha_1 ln\left(\frac{V_{it}}{X_{it}}\right) + \alpha_2 Exret_{it} + \alpha_3 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+3}\sigma_{k,it} + \Gamma Control_{it} + \varepsilon_{it} \quad (8)$$

OAS_{it} is the option adjusted spread for the respective bond as reported in the Barclays Index. An intercept is not reported as we include time fixed effects. In addition to the determinants of bankruptcy, i.e., $ln\left(\frac{V_{it}}{X_{it}}\right)$, $Exret_{it}$, $ln(E_{it})$, and $\sigma_{k,it}$, which are all issuer level determinants of credit risk, we also include issue-specific determinants of credit risk that will influence the level of credit spreads. Specifically, our additional controls include: (i) $Rating_{it}$ which is the issue-specific rating (higher rated issues are expected to have higher credit spreads, given that we code ratings to be increasing in risk), (ii) Age_{it} is the time since issuance in years (liquidity is decreasing for progressively 'off the run' securities, so we expect credit spreads to be increasing in time since issuance), and (iii) $Duration_{it}$ is option adjusted duration of the issue (for the vast majority of corporate issuers the credit term structure is upward sloping so we expect credit spreads to increase with duration, see e.g., Helwege and Turner, 1999).

For the constrained approach, we then estimate the following regression model:

$$OAS_{it} = \alpha_1 Exret_{it} + \alpha_2 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+2}\, CS_{\sigma_{k,it}} + \Gamma Control_{it} + \varepsilon_{it} \quad (9)$$

$CS_{\sigma_{k,it}}$ is the theoretical credit spread for the k^{th} measure of asset volatility. Appendix III provides a complete description of the derivation of the measures of theoretical credit spreads.

Table 1, Panel E reports the average pairwise correlations between the observed credit spread, OAS_{it}, and the theoretical credit spreads based on each volatility measure. Correlations between accounting based ($CS_{\sigma_F}, CS_{P95P5}, CS_{\sigma_F^{NAIVE}}, CS_{\sigma_F^{NAIVE(2)}}$) and market based ($CS_{\sigma_A^{NAIVE}}, CS_{\sigma_A^{\omega}}, CS_{\sigma_{AI}^{NAIVE}}, CS_{\sigma_{AI}^{\omega}}$) credit spreads are substantially higher than correlations between the corresponding volatility measures, reported in Panel D. In particular, accounting based credit spreads exhibit an average Pearson (Spearman) correlation with market based credit spreads of 0.5891 (0.4764), which contrasts with an average Pearson (Spearman) correlation of 0.2813 (0.3534) between the respective volatility measures. The reported difference in the magnitude of correlations is as expected given that theoretical credit spreads embed common additional information such as leverage.

Theoretical spreads based on historical security data or option implied volatility exhibit slightly higher correlation with observed spreads than theoretical spreads based on accounting data. In particular, OAS exhibits an average Pearson (Spearman) correlation with accounting based spreads ($CS_{\sigma_F}, CS_{P95P5}, CS_{\sigma_F^{NAIVE}}, CS_{\sigma_F^{NAIVE(2)}}$) of 0.6556 (0.5580), and an average Pearson (Spearman) correlation of 0.6591 (0.6621) with market based spreads ($CS_{\sigma_A^{NAIVE}}, CS_{\sigma_A^{\omega}}, CS_{\sigma_{AI}^{NAIVE}}, CS_{\sigma_{AI}^{\omega}}$).

3. Results

3.1 Bankruptcy forecasting

Table 2 reports the estimation results of regression equation (7). Across all specifications we find expected relations for our primary determinants: bankruptcy likelihood is decreasing in (i) distance to default barrier, $ln\left(\frac{V_{it}}{X_{it}}\right)$, (ii) recent equity returns, $Exret_{it}$, and

(iii) firm size, $ln(E_{it})$. The sample size used for the basis of estimating equation (7) is 68,104 bond-month observations. The only exception to this is when we use $\sigma_F^{NAIVE(2)}$ where the sample size drops to 60,463 bond-month observations. This reduction in sample is due to our data requirements for computing standard deviations of $RNOA$ across all fiscal quarters over the previous 20 years (requiring a minimum of 10 quarters of data).

To assess the relative importance of our different component measures of asset volatility, we first examine each measure individually after controlling for the same issuer level determinants of bankruptcy. Across models (1) to (7) in Table 2 we find that all of the component measures of asset volatility are significantly positively associated with the probability of bankruptcy. These regression specifications are unconstrained so we include each of the respective component measures of asset volatility separately and do not attempt to combine together different volatility measures. In our constrained specifications later we combine the component measures of asset volatility together.

To provide a sense of the relative economic significance across the component measures of asset volatility, we report in Panel B of Table 2 the marginal effects for each explanatory variable. Specifically, we hold each explanatory variable at its average value and report the change in probability of bankruptcy for a one standard deviation change in the respective explanatory variable relative to the full sample unconditional probability of bankruptcy. For example, column (1) in Panel B of Table 2 states that the marginal effect of σ_E is 0.0280. This means that a one standard deviation change in σ_E is associated with a 2.80% increase in bankruptcy probability, relative to the full sample unconditional probability of bankruptcy (0.62%). Comparing marginal effects across explanatory variables reveals that distance to default barrier is the most economically important explanatory variable. Individually, the most important component measure of asset volatility is σ_I (marginal effect of 0.0623 is the largest in the first 7 columns of Panel B of Table 2).

Models (8) to (12) in Table 2 combine different component measures of asset volatility. We do not include σ_E and σ_I in the same specification due to multi-collinearity (Panel D of Table 1 shows that σ_E and σ_I have a parametric correlation of 0.8816). In model (8) we start with issuer level determinants ($ln\left(\frac{V_{it}}{X_{it}}\right)$, $Exret_{it}$, and $ln(E_{it})$) and σ_I. We then add a measure of volatility from the credit markets, σ_D. Combining market based measures of asset volatility from the equity and credit markets is superior to examining equity market information alone (the pseudo-R^2 marginally increases from 29.66 percent in model (2) to 29.76 percent in model (8)). However, the coefficient on σ_D is not significant. In model (9) when we add our first measure of fundamental volatility, σ_F, we find that both σ_I and σ_F are significantly associated with bankruptcy, but σ_D is not. In terms of relative economic significance, σ_F is 1.04 times as large as that for σ_I. Using alternative measures of fundamental volatility in models (10) to (12) we find similar results: combining measures of volatility from market and accounting sources improves explanatory power of bankruptcy prediction models. In un-tabulated robustness analysis, we document further that our fundamental volatility measures also improve upon the explanatory power of a bankruptcy prediction model that includes Merton-based volatility and leverage measures (see e.g., Bharath and Shumway, 2008). This approach takes equity prices, equity volatility, and current leverage as given and then solves iteratively for asset value and asset volatility that price equity as a call option on the asset value of the firm.

One limitation with the traditional discrete hazard model analysis is that it cannot capture nonlinearities and interactions that are likely among the independent variables. As an alternative methodological approach, we analyze our default data using the Classification and Regression Trees (hereafter CART) methodology developed by Breiman, Friedman, Olshen and Stone (1984).[6] Frydman, Altman and Kao (1985) apply this technique to the prediction of financial distress and document that it outperforms discriminant analysis in out of sample tests. The data is recursively split into more homogeneous subsets, using the Gini rule to choose the optimal split at each node of the tree. Based on this approach, we generate a maximal tree and a set of sub-trees. We then use 10-fold cross validation to estimate the area under the Receiver Operating Characteristic curve (i.e., AUC) for the different sub-trees and retain the tree that maximizes the AUC. The resulting tree structure allows for non-linear and interactive associations between probability of default and the different explanatory variables, alleviating the concern that documented results are simply due to method variance.

To focus on the relative importance of accounting and market based measures of asset volatility, we first apply this technique to a basic set of bankruptcy determinants, i.e., $ln\left(\frac{V_{it}}{X_{it}}\right)$, $Exret_{it}$, $ln(E_{it})$, and a representative market-based measure of asset volatility that combines information from implied equity option data and debt market volatility, σ_{AI}^{ω}. Full details of the combined measures of asset volatility are contained in section 3.2.2. We then augment this set of variables with our fundamental volatility measures σ_F, $P95P5$, σ_F^{NAIVE} and $\sigma_F^{NAIVE(2)}$ one at a time. Panel A of Table 3 reports summary statistics for the predictive

[6] We use the Salford Predictive Modeler software suit, developed by Salford Systems, to perform the CART analysis.

ability of the resulting AUC-maximizing trees. Column (1) serves as the benchmark case when no accounting based measure of asset volatility is included. Across columns, it is clear that the test-sample (out of sample) AUC improves with the inclusion of accounting based measures of asset volatility. Note that the AUC for the augmented models ranges from 0.9165 to 0.9258, while the AUC for the basic model that only includes market volatility is 0.9057. We use bootstrap resampling to test the statistical significance of improvement in AUC. In particular, we use bootstrapping to build 100 different AUC-maximizing trees for each set of variables by changing the learn/test sample partitioning. We then compute the difference between the AUC of each of the augmented models and the AUC of the basic model. The 5th percentile of this difference is positive for the augmented models based on the two naïve asset volatility measures (columns (4) and (5)), indicating that the improvement in the AUC achieved by incorporating each of these two measures is statistically significant at conventional levels. The relative error (the simple sum of type I and type II classification errors) is also reduced by the inclusion of accounting based measures of asset volatility. In the base model the relative cost is 0.267, however with the inclusion of accounting based measures of asset volatility lowers the relative cost measure to between 0.176 and 0.218. Finally, the Hosmer-Lemehow test statistic, which takes higher values when the fit of the model is high, also increases with the inclusion of accounting based measures of asset volatility.

To further understand the economic significance of accounting based measures of asset volatility, we compute importance scores for each of the variables in the model (Panel B of Table 3). These scores are calculated as the sum of the improvement that can be attributed to a given variable at each node of the tree. Variable importance scores are then scaled between 0 and 100. Leverage is the most important variable in each of the five models. Furthermore, the importance scores of accounting based measures of asset volatility are higher than those of market based volatility measures, both considering just the role of each variable as primary splitter and its combined role as primary splitter and surrogate.

Figure 1 shows the ROC curve for each of the models. Compared to model (1), models (2) to (5), which include fundamental volatility measures, exhibit steeper ROC curves, suggesting that these models have lower type I error within the bins classified as higher risk. This evidence, together with that of Table 3, Panel A, highlights the importance of accounting based measures of asset volatility for predicting defaults out of sample.

Figure 2 presents an example of a classification tree with potential splitting variables including: $ln\left(\frac{V}{X}\right)$, $Exret$, $ln(E)$, σ_{AI}^{ω} and $\sigma_{F}^{NAIVE(2)}$. This tree has been pruned for ease of exposition. The first partition is based on $ln\left(\frac{V}{X}\right)$: high leverage firms display a higher probability of bankruptcy (4.8% vs. 0.2%). Consistent with the high importance of accounting based measures of asset volatility as a primary splitter documented in Table 3, Panel B, the second splitting variable is $\sigma_{F}^{NAIVE(2)}$. For example, within the high leverage node, firms with high fundamental volatility display a higher bankruptcy rate than firms with low fundamental volatility (5.6% vs. 0.1%). A similar split is made within the low leverage group.

Overall, the analysis in Tables 2 and 3 and Figures 1 and 2 suggest a joint role for accounting and market based measures of asset volatility in explaining corporate bankruptcies. We next turn to secondary credit markets to corroborate these results and also to better understand why accounting based measures of asset volatility would be additive to market based measures of asset volatility in explaining corporate default risk and credit spreads.

3.2 Cross-sectional variation in credit spreads

3.2.1 Unconstrained analysis

Having established the information content of our candidate component measures of asset volatility for bankruptcy prediction, we now turn to assess the information content of the same measures for secondary credit market prices. As discussed in section 2.5, under the assumption that security prices in the secondary credit market are reasonably efficient, we expect to see that the determinants of bankruptcy prediction models should also be able to explain cross-sectional variation in credit spreads.

Table 4 reports estimates of equation (8). This is our unconstrained analysis of how, and whether, different component measures of asset volatility have information content for security prices. We include month fixed effects to control for macroeconomic factors, and as such we do not report an intercept. As discussed in section 2.5, we include additional issue specific measures ($Rating_{it}$, Age_{it}, and $Duration_{it}$) to help control for other known determinants of credit spreads. Of course, it is possible that we are controlling for characteristics that subsume volatility by including these determinants, especially $Rating_{it}$. For example, the rating agencies may be using algorithms to assess credit risk that span accounting and market data sources, and as such included rating categories might subsume the ability of this data to explain cross-sectional variation in credit spreads. In unreported analysis, we find that our inferences of the combined information content of market and accounting based information to measure asset volatility are unaffected by the inclusion of $Rating_{it}$.

Across all models estimated in Table 4 we find expected relations for our primary determinants. Credit spreads are consistently decreasing in (i) distance to default barrier, $ln\left(\frac{V_{it}}{X_{it}}\right)$, and (ii) firm size, $ln(E_{it})$. Credit spreads are consistently increasing in (i) credit rating (scaled to take higher values for higher yielding issues), $Rating_{it}$, and (ii) time since issuance, Age_{it}. Recent excess equity returns, $Exret_{it}$, is usually negative across different models but is not consistently significant at conventional levels. Option adjusted duration, $Duration_{it}$, is either negatively or positively associated with credit spreads: its effect is dependent upon the included explanatory variables (once σ_D is included the relation becomes negative).

Models (1) to (7) in Table 4 examine each of our component measures of asset volatility separately. Individually, each of our component measures of asset volatility is significantly positively associated with credit spreads. To provide a sense of the relative economic significance across the component measures of asset volatility, we also report in Panel B of Table 4 the marginal effects for each explanatory variable. Similar to the marginal effects reported in Table 2, we report the change in credit spreads for a one standard deviation change for the respective explanatory variable relative to the full sample unconditional mean credit spread. Individually, the most important component measure of asset volatility is σ_I (marginal effect of 0.6181 is the largest in the first 7 columns of Panel B of Table 4).

Models (8) to (12) in Table 4 combine different component measures of asset volatility. As in Table 2, we do not include σ_E and σ_I in the same specification due to multi-collinearity concerns. In model (8) we add a measure of volatility from the credit markets,

σ_D. Consistent with the results in Table 2, combining market based measures of asset volatility from the equity and credit markets is superior to examining equity market information alone (the R^2 increases from 67.1 percent in model (2) to 73.2 percent in model (8)). In model (9) when we add our first measure of fundamental volatility, σ_F, we find that all three component measures of volatility are significantly associated with bankruptcy, but that the relative importance of σ_F is quite low. In terms of relative economic significance, σ_D is 1.05 times as large as that for σ_I, and σ_F is only 10 percent as large as that for σ_I. Using alternative measures of fundamental volatility in models (10) to (12) we find similar results: combining measures of volatility from market and accounting sources improves explanatory power of credit spreads. These results are consistent with recent research by Goodman, Neamtiu and Zhang (2013) and Sridharan (2013) who document that accounting ratios are useful in explaining returns on short dated equity options. In un-tabulated sensitivity tests, we further document that the additivity of accounting based measures of asset volatility is robust to the use of a Merton-based volatility measure estimated following Bharath and Shumway (2008).

Table 5 reports the results of equation (8) where we allow the regression coefficients to vary for Investment Grade (IG) and High Yield (HY) issuers. For the sake of brevity we only report the differential coefficients for HY issuers. As expected the HY indicator variable is strongly significantly positive reflecting the higher risk of HY issuers relative to IG issuers. Across the various specifications there is consistent evidence that the primary determinants of credit spreads are stronger for HY issuers: credit spreads are more strongly decreasing in firm size, distance to default and recent excess equity returns for HY issuers relative to IG issuers. We find that market based component measures of asset volatility, σ_E

and σ_I, and component measures of asset volatility based on fundamentals are more important for HY issuers.

3.2.2 Constrained analysis

We now assess the relative information content of the different component measures of volatility in a constrained specification. As described in Appendix III and equation (A.1), we combine component measures of asset volatility with dollar distance to default ($ln\left(\frac{V_{it}}{X_{it}}\right)$) to identify a distance to default barrier in standard deviation units. We then calibrate the various distance to default measures to an expected physical default probability which is converted to an implied spread as per equations (A.2) and (A.3). We thus generate k different theoretical spreads where the difference is attributable to the use of different component measures of asset volatility. This approach is arguably superior to the unconstrained analysis discussed in section 3.2.1 because of the inherent non-linearity between leverage, asset volatility, defaults (bankruptcy) and credit spreads. Two firms could have the same dollar distance to default but different levels of asset volatility. It is the ratio of these two measures that matters for determining physical bankruptcy probability, not the two measures separately.

An empirical challenge that we face is combining different component measures of volatility that vary in scale. As can be seen from Panel C of Table 1, the market based component measures of asset volatility have higher average values and higher standard deviations relative to the accounting based measures of asset volatility. To handle these differences in scale when we combine component measures of asset volatility we first standardize each accounting based component measure and rescale them such that they have the same mean and standard deviation as the market based component measures of asset volatility to which they will be combined with. As a result of this process we end up with seven different measures of theoretical spreads. We have four market based theoretical spreads: (i) CS_{σ_E} which is based only on historical equity volatility, (ii) CS_{σ_I} which is based on only implied equity volatility, (iii) $CS_{\sigma_A^{\omega}}$ which is based on a weighted combination of historical equity volatility and historical credit volatility, and (iv) $CS_{\sigma_{AI}^{\omega}}$ which is based on a weighted combination of implied equity volatility and historical credit volatility. We have four accounting based theoretical spreads: (i) CS_{σ_F} which is based on a parametric estimate of fundamental volatility, (ii) CS_{P95P5} which is based on a non-parametric estimate of fundamental volatility, (iii) $CS_{\sigma_F^{NAIVE}}$ which is based on historical volatility of seasonally adjusted $RNOA$ and (iv) $CS_{\sigma_F^{NAIVE(2)}}$ which is based on historical volatility of $RNOA$.

Table 6 reports regression results of equation (9). We retain the same set of controls and explanatory variables to allow comparability of explanatory power between equation (8) and equation (9). We include a set of month fixed effects and as such do not report a regression intercept. Model (1) shows that theoretical spreads based on a simple measure of historical equity volatility are able to explain 70.6 percent of the variation in credit spreads, and the regression coefficient on $CS_{\sigma_E}^{BASE}$ is 0.601. A regression coefficient less than one may suggest that our measure of theoretical credit spread is larger than the actual market spread. This is not the case as our regression model includes an intercept (via time fixed effects). In unreported analysis, if we exclude fixed effects, and other control variables, we find that the regression coefficient on $CS_{\sigma_E}^{BASE}$ is statistically greater than 1, consistent with the well-known result that structural models tend to under forecast credit spreads (e.g., Huang and Huang, 2012).

Before assessing the incremental improvement in explanatory power from alternative component measures of asset volatility, we first use our secondary credit market data to apply a 'hair-cut' to the book value of debt used as an approximation for the market value of assets. While fixed and floating rate debt is usually issued at par, changes in the credit risk of the issuer over time will create situations where the market value of debt is below the book value of debt. Thus our estimate of market value of assets may be too high (low) for issuers whose credit quality has worsened (improved) since debt issuance. A direct consequence of this is that any implied spread will be too low (high). To help mitigate this error, we take a fraction of the book value of debt as our approximation for the market value of debt using the change in the spread from when the representative bond first appears in our data set to the current time period. Specifically, we multiply the book value of debt by $\frac{1}{(1+\Delta OAS)^{Duration}}$. Thus, our estimate of the 'market' value of debt adjusts the reported book value by the change in credit spreads, ΔOAS, measured from when the representative bond was first recorded in the Barclays bond dataset to the current period. For coupon bearing debt this simply allows market value of debt to fall (rise) as credit spreads increase (decrease). Model (2) of Table 6 shows that once we incorporate this 'hair-cut' we observe a noticeable change in explanatory power. The R^2 in model (2) increases to 75.7 percent from 70.6 percent for model (1).

Models (3) to (13) in Table 6 consider various combinations of our theoretical spreads. For the sake of brevity, we do not report regression results using credit spreads based on our naïve de-levered measures of asset volatility (i.e., $CS_{\sigma_A^{NAIVE}}$ and $CS_{\sigma_{AI}^{NAIVE}}$), as we find that these are strictly dominated by credit spreads based on the weighted measures of asset volatility (i.e., $CS_{\sigma_A^{\omega}}$ and $CS_{\sigma_{AI}^{\omega}}$).

We find that the predictive power of the model is increased by combining historical (implied) equity volatility with debt volatility (the R^2 increases to 0.773 (0.774) from 0.757 (0.756), respectively). Across all market volatility based models, the model that exhibits higher R^2 is the one that includes $CS_{\sigma_{AI}^{\omega}}$. For this reason, we use this model as the basis to test for incremental information content of accounting based measures of asset volatility

Models (6) to (13) then add the four different accounting based theoretical credit spread measures. Across all four accounting based measures we see evidence of the joint role of market and accounting based component measures of asset volatility. In all specifications, accounting based volatility measures are statistically significant.

The last four rows of Table 6 contain summary information based on estimating the unconstrained regression equation (8) for the same sample of 56,846 bond-months. The sample we use in Table 6 is smaller than that in Table 4 as we require an initial out-of-sample period to empirically calibrate our distance to default to a physical bankruptcy probability. Across all of the models in Table 6 we see that the constrained regression specification results in a statistically and economically significant increase in the ability to explain cross-sectional variation in spread levels (Vuong, 1989 Z-statistics reject the null of equal explanatory power for all regression specifications). The regression specifications are identical except for how we combine leverage and volatility. The constrained specification combines leverage and volatility consistent with the Merton model, and this generates a significant improvement in explanatory power.

To help visualize the relative importance of component measures of asset volatility for credit spreads, each month we sort issuers into deciles based on $CS_{\sigma_{AI}^{\omega}}$ and CS_{σ_F}. These sorts are independent as the two measures of theoretical spreads are highly correlated (Pearson correlation of 0.7225 reported in Panel E of Table 1). We then plot the median credit spread across the resulting 100 cells. It is clear that as we move from the back to the front of Figure 3 (that is increasing theoretical spreads based on market information) we see credit spreads increase. It is also clear that as we move from left to right of Figure 3 (that is increasing theoretical spreads based on accounting information) we see also credit spreads increase. What is most interesting, though, is the increase in credit spreads along the main diagonal: when information from the market and financial statements suggest higher asset volatility credit spreads are indeed higher. A combination of market and accounting based measures of asset volatility is superior to either source alone. We find similar patterns if we instead sort issuers on the basis of $CS_{\sigma_{AI}^{\omega}}$ as an alternative market based measure of theoretical spreads, and either CS_{P95P5}, $CS_{\sigma_F^{NAIVE}}$ or $CS_{\sigma_F^{NAIVE(2)}}$ as alternative accounting based measures of theoretical spreads. For the sake of brevity we do not show these figures, but they are available upon request.

Table 7 reports the results of equation (9) where we allow the regression coefficients to vary for Investment Grade (IG) and High Yield (HY) issuers. As before in Table 5, for the sake of brevity we only report the differential coefficients for HY issuers with respect to the theoretical credit spread measures. We now find even stronger evidence that accounting based component measures of asset volatility are more relevant to explain cross-sectional variation in credit spreads for HY issuers relative to IG issuers. This inference is true for all four theoretical spreads using accounting based component measures of asset volatility: models (7), (9), (11), and (13) in Table 7 all show a statistically significant positive coefficient on the respective interaction terms.

3.3 Systematic vs. idiosyncratic volatility

The empirical analysis thus far suggests that combining market and accounting information generates superior estimates of asset volatility for forecasting bankruptcy and also for explaining cross-sectional variation in credit spreads. To help better understand the relative information content of each component measure of asset volatility we assess the extent to which market and accounting measures of returns are attributable to systematic versus idiosyncratic factors.

As discussed in the introduction, total volatility is the relevant measure of volatility for understanding credit risk, and we have priors that accounting based measures of asset volatility are more likely to be attributable to firm specific operating and investing decisions that directly impact the primary financial statements. Market based measures of asset volatility are based on changes in prices in equity and credit markets, which in turn, are driven by changes in expectations of cash flows and changes in expectations of discount rates. Arguably, the latter component is a larger determinant of changes in security prices, especially when the return interval is relatively short (e.g., Cutler, Poterba and Summers, 1989, and Richardson, Sloan and Yu, 2012). We use daily (monthly) returns as our basis for measuring equity (debt) volatility. A consequence of this is that our market based measures of asset volatility will reflect volatility of changing expectations of both cash flow news and discount rate news, with expected greater influence from the latter. In contrast, measures of volatility based on changes in accounting rates of return are a direct consequence of applying accounting rules to firm transactions over a given fiscal period. These accounting measures are mostly backward looking in terms of the cash flow generation and are only indirectly capturing changing expectations of discount rates. Our expectation for the differential impact of idiosyncratic and systematic drivers of volatility across market and accounting based measures of asset volatility also follows directly from Vuolteenaho (2002). Vuolteenaho shows that a variance decomposition of returns into cash flow news and discount rate news also maps into a firm-specific news and systematic news component. This suggests that accounting based measures of asset volatility are more likely to reflect idiosyncratic sources of volatility.

To assess the difference in the mapping of market and accounting based measures of asset volatility to systematic and idiosyncratic sources, we examine the strength of commonality across market and accounting based measures of returns. We do this by computing pairwise correlations between market and accounting based measures of returns for all possible pairs within each Fama-French sector (11 sectors in total, excluding financials). In un-tabulated analysis, we estimate these correlations using return measures over non-overlapping three-month intervals, and require at least 20 three-month periods for each pair. We use three-month returns as the shortest frequency to measure accounting based measures of returns is quarterly. There is a striking difference in the average pairwise correlation: the market-based asset volatility measure has a much higher average pairwise correlation than the accounting based measure of returns (0.43 for market based for the pooled sample and only 0.09 for accounting based). This is a necessary condition for accounting and market based return measures to differentially reflect systematic and idiosyncratic sources of risk. We also identify the first principal component for a balanced panel of 500 issuers that have non-missing credit, equity and accounting rates of return for our time period. We find that the first principal component explains 29.52 percent of the cross-sectional variation in market-based asset returns, but only 7.26 percent for accounting rates of return.[7]

The evidence suggests that market based measures of asset volatility capture relatively more systematic sources of volatility and accounting based measures of asset volatility capture more idiosyncratic sources of volatility. This provides a basis for why both market and accounting based measures are useful in generating estimates of asset volatility for forecasting bankruptcy and also for explaining cross-sectional variation in credit spreads.

3.4 Extensions and robustness tests

3.4.1 CDS data

In Table 8 we report regression estimates of a modified version of equation (9) where we use credit spreads from CDS contracts rather than bonds. As with our previous spread level regressions, we include a set of month fixed effects and as such do not report a regression intercept. A benefit of this approach is that the CDS credit spread is a cleaner representation of credit risk, but a disadvantage is the shorter time period for which this data is available (2003 to 2012 only). Because we are examining cross-sectional variation in 5 year CDS spreads, $CDS5Y_{it}$, we no longer need to control for issue specific characteristics

[7] Using a larger panel, Herskovic, Kelly, Lustig and Van Nieuwerburgh (2013) document that the first principal component explains 39% of the variation in equity returns. Further, they find that the second moments of stock return residuals from models that incorporate the Fama-French factors, exhibit high common variation.

such as Age_{it} and $Duration_{it}$. All 5 year CDS contracts have the same seniority, the same time since issuance (we only examine 'on the run' contracts), and the same tenor (5 years). Thus, we estimate the following model:

$$CDS5Y_{it} = \alpha_1 Exret_{it} + \alpha_2 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+2}\, CS_{\sigma_{k,it}} + \Gamma Control_{it} + \varepsilon_{it} \qquad (10)$$

Our sample size decreases from 56,846 bond-months examined in Table 6 to 22,199 CDS-months examined in Table 8. Despite the smaller sample size, we find striking results with this alternative sample. Models (1) to (3) show that theoretical spreads based on equity market information are able to explain up to 46.2 percent of the cross-sectional variation in credit spreads. Models (4) and (5) show that combining component measures of asset volatility generates theoretical spreads that can explain a greater fraction of the cross-sectional variation in credit spreads (the R^2 increases to 52.3 percent for model (5)). Strikingly, our measure of theoretical spread using fundamental volatility alone can explain 47.7 percent of the cross-sectional variation in credit spreads (see model (6)). Finally, including both market and accounting based measures of asset volatility yields theoretical spreads that can explain even more of the cross-sectional variation in credit spreads: a maximum R^2 of 54.1 percent across models (7), (9), (11), and (13). Similar to the analysis in Table 6, at the bottom of Table 8 we also report the R^2 of the equivalent unconstrained regression on the CDS sample. Across all of the specifications, with the exception of model (1), we see statistically significant increases in explanatory power when we constrain asset volatility and leverage, consistent with the Merton model, as compared to including these variables linearly and independently.

3.4.2 Alternative specifications

In un-tabulated analysis, we expand the bankruptcy forecasting model to control for average accounting profitability over the previous four quarters, cash holdings, market to book ratio and price, following Campbell, Hilscher and Szilagyi (2008). We choose not to include these variables in our main specification, which only includes (albeit linearly), the main determinants of probability of default as per the Merton model. Specifically, we add the following variables to the analysis reported in Table 3 (variables are defined and labelled consistently with Campbell, Hilscher and Szilagyi, 2008): (i) $NIMTAAVG$, a geometrically weighted average level of net income scaled by market value of total assets, which places higher weight on more recent quarters, (ii) $CASHMTA$, cash and short term investments scaled by the market value of assets, (iii) MB, the market to book ratio, and (iv) $PRICE$ the natural logarithm of the firm's stock price. The sample size does not change significantly as a result of the inclusion of these additional control variables (reduced from 68,104 bond-month to 67,594 bond-month observations). Our measures of fundamental volatility continue to be significant, both when included individually and together with implied volatility and debt volatility.

We also re-estimate the unconstrained and constrained credit spread regressions adding the control variables in Campbell and Taskler (2003). In particular, we control for operating income and long term debt to total assets. The sample size for the unconstrained analysis is reduced from 67,848 to 62,441 bond-month observations. Consistent with our main analysis, we continue to find that quantile-based fundamental volatility measures are significant both when included individually and when considered incrementally to debt volatility and implied volatility. Similarly, in the constrained analysis, all credit spreads based on fundamental volatility remain both individually and incrementally significant.

4. Conclusion

In this paper we evaluate alternative measures of asset volatility using information from market and accounting based sources. We find that combining accounting and market based measures of asset volatility generates superior forecasts of bankruptcy, and in turn, is better able to explain cross-sectional variation in corporate bond and corporate CDS spreads for a large sample of U.S. corporate issuers.

We further show that market based component measures of asset volatility have a greater common component to them as evidenced by greater pairwise correlations between market based measures of returns relative to accounting based measures of returns and a higher percentage of variation explained by the first principal component. This evidence suggests that market based measures reflect systematic sources of volatility and accounting based measures reflect idiosyncratic sources of volatility. Thus, combining market and accounting based measures of asset volatility generates a superior measure of total asset volatility that is relevant for understanding credit risk.

References

Altman, E (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance* 23 (4): 589-609.

Arora, N., J. Bohn and F. Zhu (2005). Reduced form vs. structural models of credit risk: a case study of 3 models. *Journal of Investment Management* 3 (4): 43-67.

Beaver, W. H. (1966). Financial ratios as predictors of bankruptcy. *Journal of Accounting Research* (4): 71-111.

Beaver, W.H., M. Correia and M. McNichols (2012). Do differences in financial reporting attributes impair the predictive ability of financial ratios for bankruptcy. *Review of Accounting Studies* 17 (4): 969-1010.

Beaver, W. H., M. McNichols, and J. Rhie (2005). Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies* 10: 93-122.

Bharath, S. and T. Shumway (2008). Forecasting default with the Merton distance to default model. *Review of Financial Studies* 21: 1339-1369.

Breiman, L., J. Friedman, C.J. Stone and R.A. Olshen (1984). Classification and regression trees. CRC Press.

Bruche, M., C. Gonzales-Aguado (2010). Recovery rates, default probabilities, and the credit cycle. Journal of Banking and Finance 34 (4): 754-764.

Campbell, J.. J. Hilscher and J. Szilagyi (2008). In search of distress risk. *The Journal of Finance* 63 (6): 2899-2939.

Campbell, J. and G. Taksler (2003). Equity volatility and corporate bond yields. *The Journal of Finance* 58 (6): 2321-2349.

Cao, C., F. Yu and Z. Zhong (2010). The information content of option-implied volatility for credit default swap valuation. *Journal of Financial Markets* 13: 321-343.

Chang, W., S. Monahan and A. Ouazad (2013). The Higher Moments of Future Return on Equity. Working Paper, Insead. http://www.insead.edu/facultyresearch/areas/accounting/events/documents/Abstract-TheHigherMomentsofFutureReturnonEquity.pdf

Chava, S., and R. Jarrow (2004). Bankruptcy prediction with industry effects. *Review of Finance* 8: 537-569.

Correia, M., S. Richardson and I. Tuna (2012). Value Investing in Credit Markets. *Review of Accounting Studies* 17 (3): 572-609.

Cremers, M., J. Driessen, P. Maenhout and D. Weinbaum (2008). Individual stock-option prices and credit spreads. *Journal of Banking & Finance* 32 (12): 2706-2715.

Crouhy, M, D. Galai and R. Mark (2000). A comparative analysis of current credit risk models. *Journal of Banking & Finance* 24: 59-117.

Cutler, D., J. Poterba and L. Summers (1989). What moves stock prices? *The Journal of Portfolio Management* 15 (3): 4-12.

Fama, E., and K. French (2000). Forecasting profitability and earnings. *Journal of Business* 72: 161–175.

Feldhutter, P., and S. Schaefer (2013). The credit spread puzzle – myth or reality? Working paper, London Business School. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2363081

Frydman, H., E. Altman, and D. Kao (1985). Introducing recursive partitioning for financial classification: the case of financial distress. *Journal of Finance* 40 (1): 269-291.

Gemmill, G. and A. Keswani (2011). Downside risk and the size of credit spreads. *Journal of Banking and Finance* 35 (8): 2021-2036.

Goodman, T., M. Neamtiu and F. Zhang. (2012). Fundamental Analysis and Option Returns. Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1974753

Haesen D., P. Howeling and V. Van Zundert (2012). Equity Momentum for Corporate Bonds: Identifying, Understanding and Reducing Time-Varying Factor Exposures. Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2131032

Helwege, J., and C. M. Turner (1999). The slope of the credit yield curve for speculative grade issuers. *Journal of Finance* 54 (5), 1869-1884.

Herskovic, B., B. Kelly, H, Lustig, and S.Van Nieuwerburgh (2013) The Common Factor in Idiosyncratic Volatility. Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2174541

Hou, K., M.A. Van Dijk and Y. Zhang (2012). The Implied Cost of Capital: A New Approach. *Journal of Accounting and Economics* 53(3): 504-526.

Huang, J. and M. Huang (2012). How much of the corporate-treasury yield spread is due to credit risk? *Review of Asset Pricing Studies* v 2 (2): 153-202.

Kealhoffer, S.(2003). Quantifying Credit Risk II: Debt Valuation. *Financial Analysts Journal*, May/June: 78-92.

Konstantinidi, T. and P. Pope (2012). Forecasting Risk in Earnings. Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1903378.

Lok, S. and S. Richardson (2011). Credit markets and financial information. *Review of Accounting Studies* 16: 487-500.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance* 29: 449-470.

Ohlson, J. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* 18 (1): 109-131

Penman, S. H. (1991). An Evaluation of Accounting Rate-of-Return. *Journal of Accounting, Auditing and Finance* Spring: 233-255.

Penman, S. H. (2014). Financial statement analysis and security valuation. 5th edition. McGraw-Hill International Edition.

Richardson, S. A., R. G. Sloan, M. T. Soliman and I. Tuna (2006). The implications of accounting distortions and growth for accruals and profitability. *The Accounting Review* 81 (3): 713-743.

Richardson, S.A., R.G. Sloan and H. You (2012). What makes stock prices move? Fundamentals vs. investor recognition. *Financial Analysts Journal* 68 (2).

Schaefer, S. and I. Strebulaev (2008). Structural models of credit risk are useful: evidence from hedge ratios on corporate bonds. *Journal of Financial Economics* 90 (1): 1-19.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business* 74: 101-124.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings?. *The Accounting Review* 81 (3): 289-315.

Sridharan, S. (2013). Volatility Forecasting using Fundamental Information. Working Paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1984324

Vuolteenaho, T. (2002). What drives firm-level stock returns? *The Journal of Finance* 57 (1): 233-264.

Vuong, Q. H. (1989). Likelihood ratio-tests for model selection and non-nested hypotheses. *Econometrica* 57 (2): 307-333.

Appendix I: Variable Definitions
Compustat/CRSP mnemonics in parenthesis

Panel A: Volatility measures

Variable	Description
σ_E	Historical equity volatility, the annualized standard deviation of realized daily stock returns over the previous 252 days.
σ_I	Implied volatility, the average of implied Black and Scholes volatility estimates for at-the-money 91-day call and put options (source: Option Metrics Ivy DB standardized database).
σ_D	Debt volatility, the annualized standard deviation of total monthly bond returns, computed over the previous 12 months (computed based on Barclays Capital total return).
σ_A^{NAIVE}	Naively deleveraged historical equity volatility, $\sigma_E\omega$, where ω is defined as in Panel B.
σ_A^{ω}	Weighted historical volatility, $\sqrt{\omega^2\sigma_E^2 + (1-\omega)^2\sigma_D^2 + 2\omega(1-\omega)\rho_{D,E}\sigma_E\sigma_D}$, where ω and $\rho_{D,E}$ are defined as in Panel B.
σ_{AI}^{NAIVE}	Naively deleveraged implied equity volatility, $\sigma_I\omega$, where ω is defined as in Panel B.
σ_{AI}^{ω}	Weighted implied volatility, $\sqrt{\omega^2\sigma_I^2 + (1-\omega)^2\sigma_D^2 + 2\omega(1-\omega)\rho_{D,E}\sigma_I\sigma_D}$, where ω and $\rho_{D,E}$ are defined as in Panel B.
σ_F	Fundamental volatility, the standard deviation of the estimated $RNOA$ percentiles ($RNOA$ is computed for each quarter as the rolling sum of 'OIADP' for the previous 4 quarters, scaled by the average of the opening and ending balance of NOA over this 4 quarter period).
$P95P5$	The difference between the estimated 95th and 5th percentiles of the $RNOA$ distribution.
σ_F^{NAIVE}	Standard deviation of the difference between quarterly $RNOA$ and $RNOA$ for the same quarter of the previous year, computed over the previous 5 years (requiring a minimum of 10 quarters of data).
$\sigma_F^{NAIVE(2)}$	Average standard deviation of quarterly $RNOA$. The standard deviations of $RNOA$ for fiscal quarters 1, 2, 3 and 4 are computed over the previous 20 years (requiring a minimum of 10 quarters of data). The resulting quarter-specific volatilities are then averaged across the four fiscal quarters.

Panel B: Credit spreads and other variables used in the estimation of asset volatility and theoretical credit spreads

Variable	Description
OAS	Option adjusted spread (source: Barclays Capital bond data).
$Duration$	Option adjusted duration (source: Barclays Capital bond data).
Age	Number of years from the date of issuance to the current month, calculated as (current date-issue date)/365.
$Rating$	Barclays Capital index rating, converted to a numeric scale. Rating ranges from 1 (index rating AAA) to 21 (index rating C).
$Exret$	Excess returns, the difference between equity returns and value weighted market returns over the last 12 months.
STD	Book value of short term debt (‘DLCQ’).
LTD	Book value of long term debt (‘DLTTQ’).
X	Book value of short term debt (STD)+0.5* book value of long term debt (LTD).
E	Market capitalization, calculated as |‘PRC’|*’SHROUT’/1,000. For firms with multiple classes of shares, we add the market value of each class of shares (source: CRSP monthly file)
ω	$\frac{E}{E+STD+LTD}$, market capitalization scaled by the sum of market capitalization and the book value of debt (where book value of debt is defined as $STD + LTD$).
$r_{i,t}^2$	Correlation between the firm’s monthly equity return and the market value weighted return calculated over the prior 5 years (computed based on the CRSP monthly file).
$\rho_{E,D}$	Average correlation of monthly equity and bond returns, calculated over the prior 12 months for all bonds in the same decile of OAS (computed based on the equity returns from the CRSP monthly file and total bond returns from Barcap).
μ	The drift in asset value, defined as $\mu = r_f + \beta RP$, where r_f is the one-year swap rate, available at St. Louis Fed website, RP is the market risk premium, which we set equal to 4%, and β is the asset beta of the firm. β is defined as the coefficient from the rolling regression of the firm’s monthly asset returns over the previous 24 months on the average asset returns calculated accross all firms, requiring at least 12 months of available data. We compute asset returns by weighting the respective equity and credit return each month by the respective weight of equity (ω) and credit $(1 - \omega)$ in the capital structure of the firm.

Panel B (Cont.)

Variable	Description
δ	The payout ratio, calculated as the sum of interest payments to debtholders over the previous four quarters (calculated using 'INTPNY'), the dividend payments to equityholders (the product of the annual dividend 'DVI' and the number of shares outstanding, 'CSHOC', both obtained from the 'Security daily' module of Compustat/CRSP merged database) and purchases of common and preferred stock over the previous four quarters (calculated using 'PRSTKCY'), scaled by the firm's total assets ($E + STD + LTD$).
V	Sum of the market capitalization of equity plus and the book value of short term debt (STD) and long term debt (LTD).
V^{Alt}	$E + \frac{STD+LTD}{(1+\Delta OAS)^{Duration}}$, where ΔOAS is the difference between the current option adjusted spread (OAS) and the option adjusted spread for the first month the bond is in the sample.

Panel C: Fundamental volatility estimation

Variable	Description
$RNOA$	Return on net operating assets, defined as operating income after depreciation ('OIADP') scaled by average of the opening and closing balance of net operating assets (NOA).
NOA	Net operating assets, defined as the sum of common equity, preferred stock, long-term debt, debt in current liabilities and minority interests minus cash and short term investments, 'CEQ'+'PSTK'+'DLTT'+'DLC'+'MIB'-'CHE'.
ACC	Accruals scaled by the average of the opening and closing balance of NOA, with accruals calculated as Δ'ACT'-Δ'CHE'-(Δ'LCT'- Δ'DLC'- Δ 'TXP')-'DP', where 'ACT' are current assets, 'CHE' cash and short term investments, 'LCT' current liabilities, 'DLC' debt in current liabilities, 'TXP' taxes payable and 'DP' depreciation and amortization.
$LOSS$	An indicator variable equal to 1 if $RNOA$<0, 0 otherwise.
$PAYER$	An indicator variable equal to 1 if $Payout$>0, 0 otherwise.
$PAYOUT$	Dividends paid, 'DVPSX_F', scaled by the average opening and closing balances of $RNOA$.

Panel D: Credit spreads

Variable	Description
$CS_{\sigma_E}^{BASE}$	$CS_{\sigma_E}^{BASE} = -\frac{1}{T}[1-(1-R)CQDF]$, where $CQDF = N\left[N^{-1}(CPD) + \lambda\sqrt{r^2}\sqrt{T}\right]$ and $CPD = 1-(1-PD)^T$ and PD is the empirically fitted physical probability of default, resulting from the estimation of the following logistic regression $E(PD) = f\left(\frac{ln\frac{V}{X}+\left(\mu-\delta-\frac{\sigma_E^2}{2}\right)t}{\sigma_E\sqrt{t}}\right)$. Please refer to Appendix III for more details on the calculation of theoretical credit spreads.
CS_{σ_K}	Similar to $CS_{\sigma_E}^{BASE}$, except that $E(PD) = f\left(\frac{ln\frac{V^{Alt}}{X}+\left(\mu-\delta-\frac{\sigma_k^2}{2}\right)t}{\sigma_k\sqrt{t}}\right)$, where σ_k are the different measures of volatility described in Panel A, t is the option adjusted duration and the remaining parameters are defined as in Panel B. Please refer to Appendix III for more details on the calculation of theoretical credit spreads.

Appendix II: Quantile regression approach

In this Appendix we describe the quantile regression approach discussed in Section 2.3.3.2. We use this approach to estimate the quantiles and conditional moments of the $RNOA$ distribution. For each year t, we estimate equation (6) using quarterly data from 1963 to t. Our model is similar to the one in Chang, Monahan and Ouazad (2013) and Hou, Van Dijk, and Zhang (2012), with the exception that we forecast return on net operating assets ($RNOA$) instead of return on equity (ROE) and therefore do not include leverage as an explanatory variable and scale all variables by the average balance of net operating assets (NOA) rather than by the average balance of book equity. All variables used in the estimation are described in Appendix I. We compute these variables at the end of each quarter, using the most recent four quarters of data.

In unreported analyses, we find the expected relations between our included explanatory variables and future profitability. Specifically, the median quantile regression generates the following results: (i) β_1^{50} is 0.94 consistent with mean reversion in accounting rates of return (e.g., Penman, 1991, and Fama and French, 2000), (ii) β_2^{50} is -0.01 consistent with loss makers having lower levels of future profitability (e.g., Hou, Van Dijk, and Zhang, 2012), (iii) β_3^{50} is -0.14 consistent with faster mean reversion in profitability for loss making firms (e.g., Beaver, Correia and McNichols, 2012), (iv) β_4^{50} is -0.02 consistent with the well documented negative relation between accruals and future firm performance (e.g., Sloan, 1996, and Richardson, Sloan, Soliman and Tuna, 2006), (v) β_5^{50} is 0.02 consistent with dividend paying firms having higher levels of future profitability (e.g., Hou, Van Dijk, and

Zhang, 2012), and (vi) β_6^{50} is 0.26 also consistent with firms with higher dividend payout having higher levels of profitability (e.g., Hou, Van Dijk, and Zhang, 2012).

We combine the values of the independent variables in year t with the vector of coefficients, $\text{Ɓ}_t^q = \beta_{0t}^q, \ldots, \beta_{6t}^q$, to obtain out-of-sample estimates of the percentiles for the year t+1. In particular, we obtain a vector of coefficient estimates, $\widehat{\text{Ɓ}_t^q}$, for each percentile and sample quarter. Based on this vector, we estimate the expected value of each of the 100 percentiles as $E(\widehat{q_{\iota t+1}}|X_{\iota t}) = \widehat{\text{Ɓ}_t^q} X_{it}$, where X_{it} includes $RNOA_{it}$, $LOSS_{it}$, $LOSS_{it} \times RNOA_{it}$, ACC_{it}, $PAYER_{it}$, $PAYOUT_{it}$.

For purposes of estimation of the vector of coefficient estimates, we delete extreme observations of dependent and independent variables. In particular, we delete all observations with $|RNOA_{it}|$>2, $|RNOA_{it-1}|$>2, $|ACC_{it-1}|$>2, $|PAYOUT_{it-1}|$>1, $|PAYOUT_{it-1}|$<0. We retain all values of these variables, irrespective of extreme values, when we generate the expected quantile values.

We focus on two measures of conditional volatility for each firm, and year t. First, we define the standard deviation of the distribution of quantile estimates, $\sigma_F = Std\big(E(\widehat{q_{\iota t+1}}|X_{\iota t})\big)$, q=1,…, 100. Second, we define the difference between the predicted value of the 95th percentile and the predicted value of the 5th percentile, $P95P5 = E(\widehat{95_{\iota t+1}}|X_{\iota t}) - E(\widehat{5_{\iota t+1}}|X_{\iota t})$.

Appendix III: Theoretical Credit Spreads

In this Appendix we describe the calculation of theoretical credit spreads. We first combine our measures of the dollar distance to default, $ln\left(\frac{V_{it}}{X_{it}}\right)$, and the respective measures of asset volatility, $\sigma_{k,it}$, to construct a measure of expected distance to default. The expected distance to default measure also includes a drift term $\left(\mu_{it} - \delta_{it} - \frac{\sigma_{A,it}^2}{2}\right)t$. μ is defined as $r_f + \beta RP$, where r_f is the 1-year swap rate, RP is the market risk premium, which we set equal to 4% and β is the asset beta of the firm, the coefficient from a rolling regression of the firm's monthly asset returns over the previous 24 months on the average asset returns, requiring at least 12 months of available data. Following Feldhutter and Schaefer (2013), the payout ratio, δ, is calculated as the sum of interest payments to debt-holders over the previous four quarters (based on 'INTNY'), the dividend payments to equity-holders (the product of the annual dividend 'DVI' and the number of shares outstanding 'CSHOC') and purchases of common and preferred stock over the previous four quarters (based on 'PRSTKCY'), scaled by the firm's total assets ($E + STD + LTD$). This distance to default is then empirically mapped to our bankruptcy data using a discrete time hazard model to generate a forecast of physical bankruptcy probability, labelled as $E\left(PD_{it}^k\right)$. We estimate this physical bankruptcy probability for each of our asset volatility measures according to equation (A.1) below:

$$E\left(PD_{it}^k\right) = f\left[\frac{ln\frac{V_{i0}}{X_{it}} + \left(\mu_{it} - \delta_{it} - \frac{\sigma_{A,it}^2}{2}\right)t}{\sigma_{A,it}\sqrt{t}}\right] \quad \text{(A.1)}$$

We next convert each physical bankruptcy probability into a risk-neutral measure, following the approach described in Kealhofer (2003) and Arora, Bohn, and Zhu (2005). We first compute the cumulative physical bankruptcy probability, CPD_{it}^{k}, from $E(PD_{it}^{k})$ by cumulating survival probabilities over the relevant number of periods. In particular, $CPD_{it}^{k} = 1 - \left(1 - E(PD_{it}^{k})\right)^{T}$. We then convert this cumulative physical bankruptcy probability, CPD_{it}^{k}, to a cumulative risk neutral bankruptcy probability, $CQDF_{it}^{k}$. We use a normal distribution to convert physical probabilities of bankruptcy to risk neutral probabilities, following the approach in Crouhy, Galai, and Mark (2000), Kealhofer (2003), and Arora, Bohn and Zhu (2005):

$$CQDF_{it}^{k} = N\left[N^{-1}\left[CPD_{it}^{k}\right] + \lambda\sqrt{r_{it}^{2}}\sqrt{T}\right] \quad \text{(A.2)}$$

The cumulative physical bankruptcy probability is first converted into a point in the cumulative normal distribution. A risk premium is then added. The risk premium is the product of (i) the issuer's sensitivity to the market price of risk, as measured by the correlation between the underlying issuer-level asset returns and the market index return, $\sqrt{r_{it}^{2}}$, (ii) the market price of risk (i.e. the market Sharpe ratio, measured by λ), and (iii) the duration of the credit risk exposure, T. The risk modified physical bankruptcy probability is then mapped back to risk neutral space. We set the market Sharpe ratio, λ, equal to 0.5, consistent with the values observed by Kealhofer (2003). We set $\sqrt{r_{it}^{2}}$ equal to the correlation between monthly firm stock returns and monthly market returns using a rolling 60-month window. We impose a floor (ceiling) on the estimated correlation at 0.1 (0.7). Finally, we estimate implied (or theoretical) credit spreads as follows:

$$CS_{it}^{k} = -\frac{1}{T} ln\left[1 - (1 - R_{it})CQDF_{it}^{k}\right] \quad \text{(A.3)}$$

R_{it} is expected recovery rate conditional on bankruptcy, which we set equal to 0.4 for all firms. While we assume R_{it} to be a constant, it is possible that recovery rates exhibit systematic time-variation (Bruche and Gonzales-Aguado, 2010). While this could affect the gap between theoretical and observed credit spreads, we have no reason to believe it will present a concern to our analysis, given that we do not examine this gap directly.

Figure 1
Receiver Operating Curve
Binary Recursive Partitioning Analysis

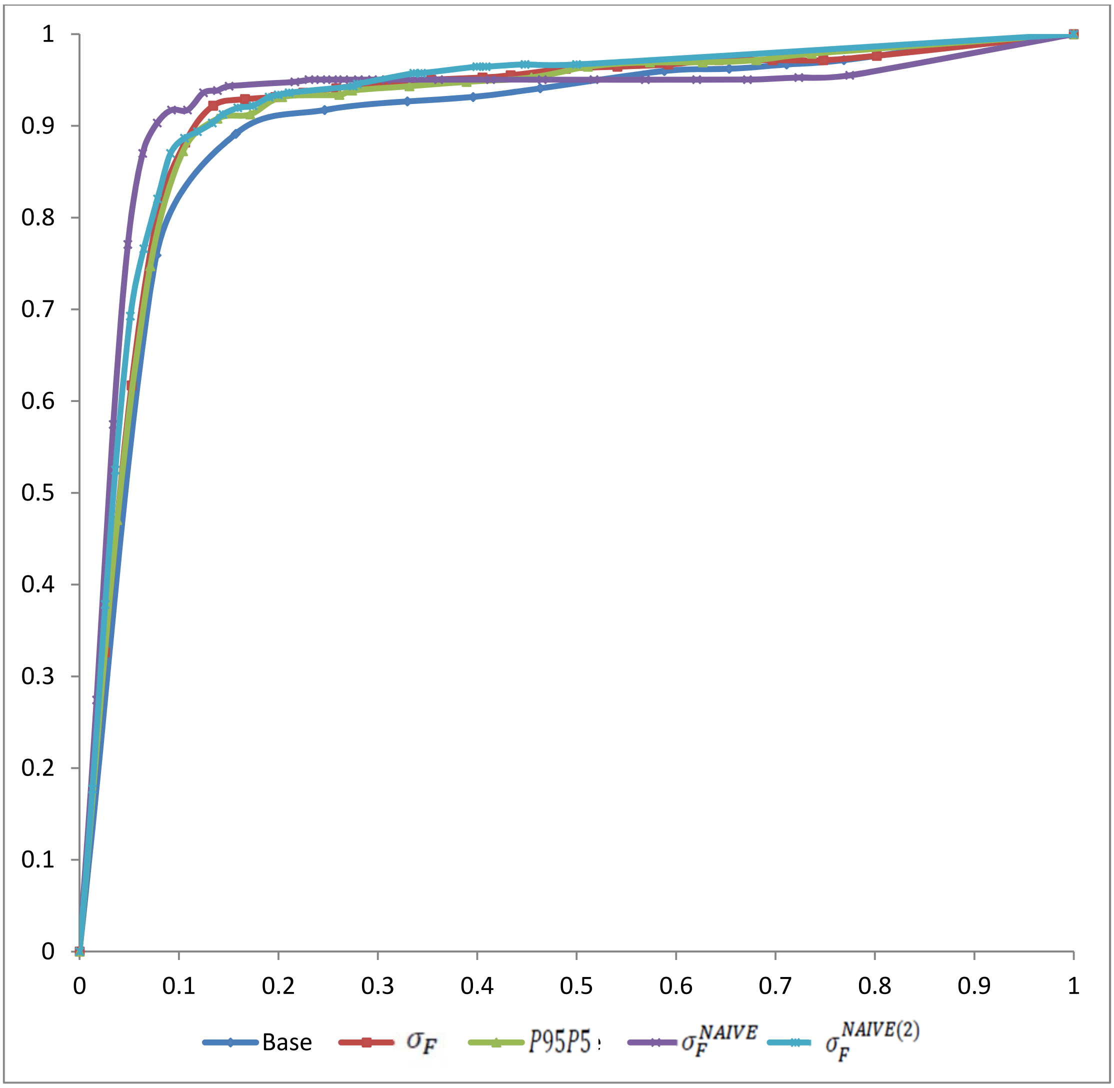


Figure 2: Binary Recursive Partitioning Example

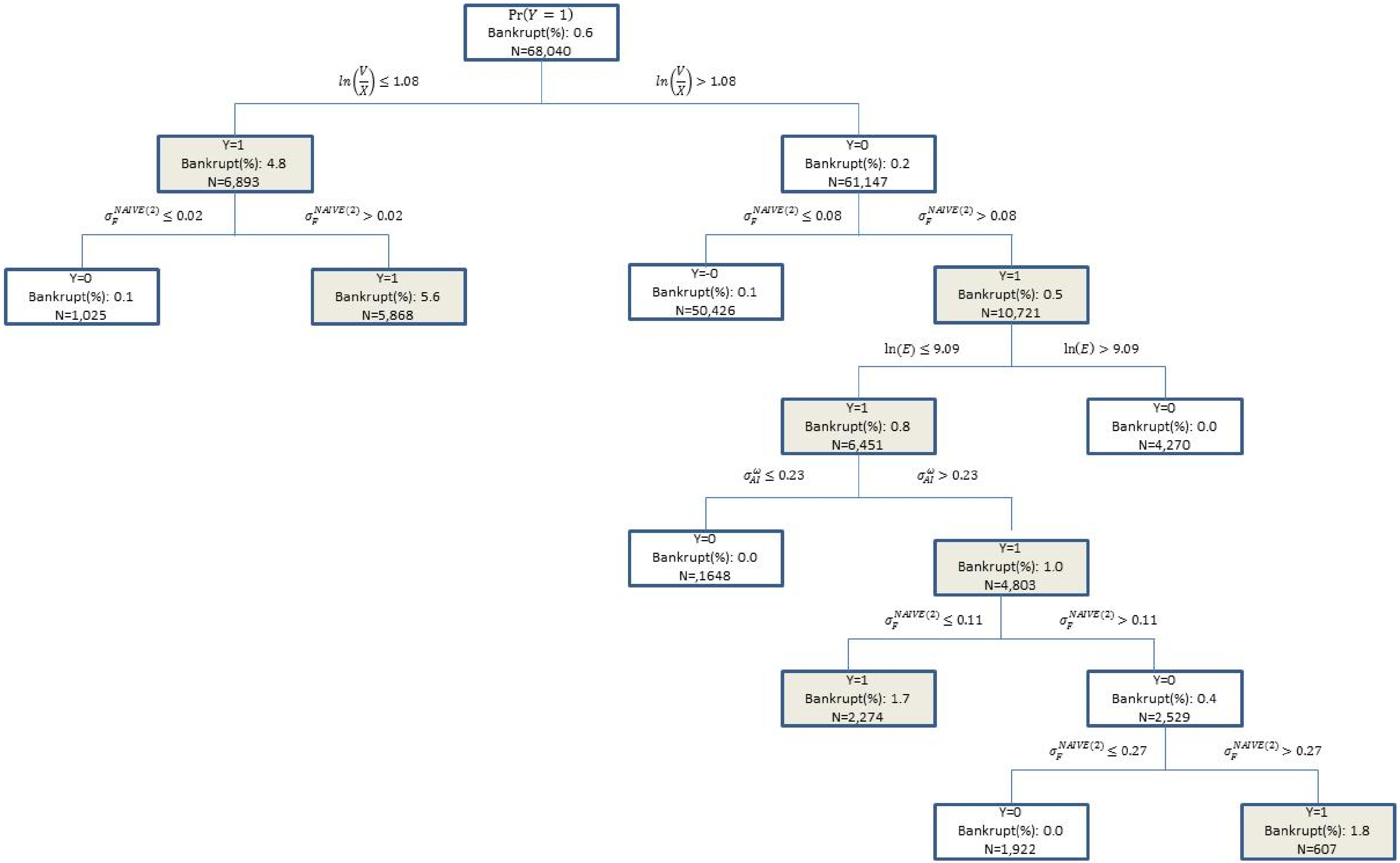
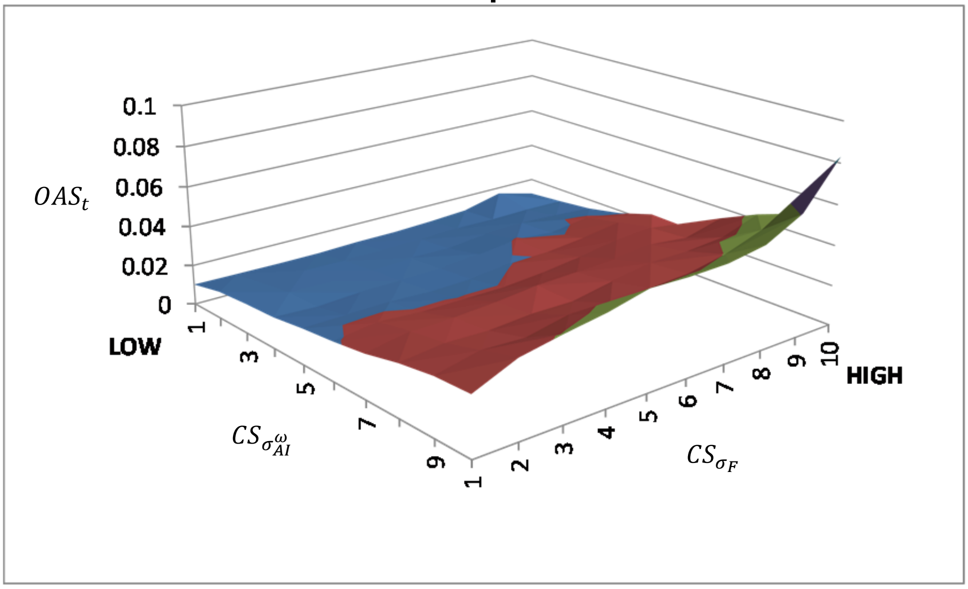


Figure 3
Median credit spreads by decile of market based and fundamental based theoretical credit spreads



Each month we sort issuers into deciles based on $CS_{\sigma_{AI}^{\omega}}$ and CS_{σ_F}. These sorts are independent given that our sorting variables are highly correlated. We then plot the median credit spread across the resulting 100 cells.

Table 1
Descriptive statistics

Panel A: Industry composition

	%
Consumer Non-Cyclical	17.36
Consumer Cyclical	17.22
Capital Goods	10.42
Basic Industry	10.16
Energy	9.77
Communications	9.03
Electric	8.14
Technology	6.36
Other Industrial	4.46
Transportation	3.39
Natural Gas	2.53
Other	1.17

Panel B: Bond characteristics

	N	Mean	Std. Dev.	p1	p25	Median	p75	p99
OAS	121,270	0.0331	0.0513	0.0000	0.0096	0.0194	0.0400	0.2421
$Duration$	121,257	5.1602	2.2016	0.7200	4.0300	5.0000	5.9600	12.5900
Age	119,438	2.8921	2.4375	0.1178	1.1589	2.3986	3.9863	12.7342
$Rating$	120,862	10.3637	4.0371	2.0000	7.0000	10.0000	14.0000	19.0000
$Exret$	120,604	0.0020	0.1267	-0.3293	-0.0537	-0.0011	0.0525	0.3719
$ln\left(\frac{V}{X}\right)$	121,224	1.8239	0.7326	0.6336	1.2785	1.7092	2.2561	3.9473
$ln(E)$	121,270	7.9814	1.6970	3.6800	6.9011	8.0109	9.1292	11.8345
ω	121,270	0.6348	0.2212	0.0652	0.4944	0.6691	0.8114	0.9712
$r_{i,t}^2$	121,130	0.2063	0.1566	0.0005	0.0803	0.1764	0.3026	0.6276
$\rho_{E,D}$	121,270	0.2195	0.1511	0.0500	0.0729	0.1894	0.3403	0.5708

Panel C: Volatility measures

	N	Mean	Std. Dev.	p1	p25	Median	p75	p99
σ_E	120,007	0.4082	0.2371	0.1312	0.2497	0.3446	0.4885	1.2953
σ_I	92,902	0.3903	0.1908	0.1434	0.2603	0.3446	0.4648	1.0850
σ_D	92,638	0.0887	0.1048	0.0140	0.0428	0.0582	0.0875	0.6045
σ_A^{NAIVE}	120,007	0.2338	0.1215	0.0500	0.1519	0.2115	0.2892	0.6588
σ_A^{ω}	92,143	0.2562	0.1375	0.0781	0.1663	0.2250	0.3074	0.7784
σ_{AI}^{NAIVE}	92,902	0.2432	0.1098	0.0676	0.1689	0.2269	0.2967	0.6041
σ_{AI}^{ω}	71,381	0.2581	0.1177	0.0820	0.1794	0.2366	0.3099	0.6758
σ_F	117,893	0.0528	0.0393	0.0137	0.0319	0.0447	0.0595	0.2437
$P95P5$	117,893	0.1739	0.1341	0.0418	0.1041	0.1467	0.1953	0.8470
σ_F^{NAIVE}	114,323	0.0586	0.1092	0.0056	0.0174	0.0306	0.0558	0.6954
$\sigma_F^{NAIVE(2)}$	100,289	0.0662	0.0914	0.0102	0.0270	0.0424	0.0688	0.5752

Panel D: Correlations across volatility measures

	σ_E	σ_I	σ_D	σ_A^{NAIVE}	σ_{AI}^{NAIVE}	σ_A^{ω}	σ_{AI}^{ω}	σ_F	$P95P5$	σ_F^{NAIVE}	$\sigma_F^{NAIVE(2)}$
σ_E	1	0.8816	0.3813	0.5057	0.7113	0.4626	0.6760	0.1019	0.0856	0.1811	0.1974
σ_I	0.9008	1	0.4874	0.4021	0.6011	0.4912	0.7023	0.1625	0.1435	0.1929	0.1795
σ_D	0.2651	0.3371	1	-0.0708	0.2642	-0.0175	0.2978	0.0812	0.0682	0.0868	0.0515
σ_A^{NAIVE}	0.5468	0.4723	-0.0591	1	0.8723	0.9150	0.8184	0.2627	0.2649	0.2254	0.2705
σ_{AI}^{NAIVE}	0.7102	0.6275	0.1460	0.8912	1	0.8265	0.9169	0.3023	0.2985	0.2396	0.2676
σ_A^{ω}	0.5031	0.5500	-0.0110	0.9180	0.8438	1	0.8871	0.3056	0.3043	0.2759	0.2764
σ_{AI}^{ω}	0.6658	0.7168	0.1649	0.8425	0.9174	0.9166	1	0.3286	0.3216	0.2799	0.2767
σ_F	0.0180	0.0792	0.0004	0.2754	0.2858	0.3200	0.3164	1	0.9979	0.4206	0.4048
$P95P5$	0.0016	0.0587	-0.0127	0.2818	0.2868	0.3235	0.3137	0.9977	1	0.4201	0.4040
σ_F^{NAIVE}	0.3530	0.3524	0.0839	0.3581	0.3958	0.3897	0.4177	0.2115	0.2080	1	0.6356
$\sigma_F^{NAIVE(2)}$	0.3845	0.3506	0.0654	0.4201	0.4217	0.4142	0.4339	0.2667	0.2646	0.6323	1

Panel E: Correlations between actual and implied credit spreads

	OAS	$CS_{\sigma_E}^{BASE}$	CS_{σ_E}	CS_{σ_i}	$CS_{\sigma_A^{NAIVE}}$	$CS_{\sigma_{Ai}^{NAIVE}}$	$CS_{\sigma_A^{\omega}}$	$CS_{\sigma_{AI}^{\omega}}$	CS_{σ_F}	CS_{P95P5}	$CS_{\sigma_F^{NAIVE}}$	$CS_{\sigma_F^{NAIVE(2)}}$
OAS	1	0.7360	0.7731	0.7450	0.5162	0.5784	0.7714	0.7702	0.6693	0.6648	0.6581	0.6303
$CS_{\sigma_E}^{BASE}$	0.7759	1	0.9776	0.9182	0.5200	0.5401	0.8473	0.8355	0.7410	0.7359	0.7223	0.7164
CS_{σ_E}	0.7824	0.9988	1	0.9480	0.4852	0.5171	0.8678	0.8622	0.7768	0.7719	0.7611	0.7538
CS_{σ_i}	0.7424	0.9607	0.9631	1	0.5169	0.6016	0.8455	0.8983	0.7571	0.7509	0.7682	0.7660
$CS_{\sigma_A^{NAIVE}}$	0.5813	0.7614	0.7476	0.7519	1	0.8837	0.6705	0.6888	0.4466	0.4433	0.4526	0.4465
$CS_{\sigma_{Ai}^{NAIVE}}$	0.6214	0.7791	0.7726	0.8239	0.9172	1	0.6779	0.7663	0.4693	0.4645	0.4633	0.4534
$CS_{\sigma_A^{\omega}}$	0.7350	0.9024	0.8974	0.8505	0.9262	0.8968	1	0.9435	0.7484	0.7450	0.7323	0.7012
$CS_{\sigma_{AI}^{\omega}}$	0.7105	0.8680	0.8652	0.9087	0.8910	0.9632	0.9378	1	0.7225	0.7177	0.7257	0.6939
CS_{σ_F}	0.5329	0.6734	0.6826	0.6285	0.3736	0.3890	0.5142	0.4760	1	0.9981	0.8254	0.7814
CS_{P95P5}	0.5246	0.6704	0.6798	0.6257	0.3672	0.3826	0.5080	0.4698	0.9988	1	0.8284	0.7862
$CS_{\sigma_F^{NAIVE}}$	0.5971	0.7810	0.7915	0.7710	0.4394	0.4794	0.5853	0.5747	0.7937	0.7971	1	0.8667
$CS_{\sigma_F^{NAIVE(2)}}$	0.5774	0.7593	0.7693	0.7540	0.4427	0.4799	0.5714	0.5685	0.7767	0.7798	0.9259	1

Correlations are computed for each of the months for which we have data. Correlations are based on the largest possible sample size for each pair of default forecasts. Reported correlations are averages across the months in the sample. Average Pearson correlations are reported above the diagonal and average Spearman correlations are reported below the diagonal. Variable definitions are provided in Appendix I.

Table 2
Probability of Bankruptcy

$$\Pr(Y_{it+1} = 1) = f\left[ln\left(\frac{V_{it}}{X_{it}}\right), Exret_{it}, ln(E_{it}), \sigma_{k,it}\right] \quad (7)$$

Panel A: Regression analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Intercept	1.385	-0.680	2.177**	2.269***	2.833***	2.196***	1.951**	-0.715	-0.608	-0.583	-0.558	-0.769
	(1.16)	(-0.52)	(2.48)	(2.60)	(3.30)	(2.66)	(2.02)	(-0.55)	(-0.46)	(-0.44)	(-0.42)	(-0.58)
$ln\left(\frac{V}{X}\right)$	-2.377***	-2.025**	-2.771***	-2.783***	-2.780***	-2.937***	-2.260***	-1.955**	-2.229**	-2.231**	-2.168**	-2.483***
	(-2.94)	(-2.51)	(-3.47)	(-3.47)	(-3.35)	(-3.29)	(-2.77)	(-2.37)	(-2.56)	(-2.56)	(-2.46)	(-2.63)
$Exret$	-1.224**	-0.655*	-0.983*	-0.988*	-0.960*	-1.421**	-1.307**	-0.767**	-0.727*	-0.727*	-0.698*	-0.906**
	(-2.36)	(-1.69)	(-1.83)	(-1.83)	(-1.71)	(-2.17)	(-2.52)	(-2.06)	(-1.91)	(-1.91)	(-1.90)	(-1.99)
$ln(E)$	-0.526***	-0.400***	-0.575***	-0.579***	-0.601***	-0.513***	-0.579***	-0.404***	-0.404***	-0.403***	-0.393***	-0.326**
	(-4.17)	(-3.15)	(-4.29)	(-4.32)	(-4.55)	(-3.77)	(-4.65)	(-3.15)	(-3.03)	(-3.04)	(-3.01)	(-2.44)
σ_E	1.120***											
	(2.60)											
σ_I		2.381***						2.162***	1.768***	1.792***	2.011***	2.099***
		(4.66)						(4.31)	(3.49)	(3.53)	(4.02)	(3.98)
σ_F			11.708***						8.959***			
			(5.83)						(3.46)			
$P95P5$				3.347***						2.548***		
				(5.69)						(3.38)		
σ_F^{NAIVE}					2.731***						1.985**	
					(3.67)						(2.30)	
$\sigma_F^{NAIVE(2)}$						3.691***						3.379***
						(3.04)						(2.74)
σ_D							1.944***	0.643	0.273	0.288	0.425	-0.469
							(2.99)	(0.97)	(0.37)	(0.40)	(0.61)	(-0.55)
Nobs	68,104	68,104	68,104	68,104	68,104	60,463	68,104	68,104	68,104	68,104	68,104	60,463
Pseudo-R2	0.2745	0.2966	0.2964	0.295	0.2816	0.2725	0.2780	0.2976	0.3133	0.3125	0.3058	0.2913

Panel B: Marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)	*(7)*	*(8)*	(9)	(10)	(11)	(12)
	Effect of a one standard deviation change on the probability of bankruptcy scaled by the unconditional probability of bankruptcy one year ahead											
$ln\left(\frac{V}{X}\right)$	-0.2090	-0.2064	-0.1984	-0.2000	-0.1929	-0.1627	-0.2125	-0.2072	-0.2072	-0.2088	-0.2030	-0.1734
$Exret$	-0.0166	-0.0103	-0.0109	-0.0110	-0.0103	-0.0122	-0.0190	-0.0125	-0.0104	-0.0105	-0.0101	-0.0098
$ln(E)$	-0.0927	-0.0817	-0.0826	-0.0834	-0.0837	-0.0569	-0.1091	-0.0857	-0.0752	-0.0756	-0.0739	-0.0456
σ_E	0.0280											
σ_I		0.0623						0.0589	0.0422	0.0431	0.0484	0.0377
σ_F			0.0441						0.0438			
$P95P5$				0.0432						0.0428		
σ_F^{NAIVE}					0.0282						0.0277	
$\sigma_F^{NAIVE(2)}$						0.0249						0.0287
σ_D							0.0219	0.0082	0.0031	0.0032	0.0048	-0.0039

Variable definitions are provided in Appendix I. Standard errors are clustered by firm and month. Regressions are based on a sample of 68,104 firm-months for the period January 1996 through to December 2012.
Marginal effects are reported as the marginal increase in the probability of bankruptcy as each of the explanatory variables increases by one standard deviation, scaled by the unconditional probability of bankruptcy one year ahead.

Table 3
Binary Recursive Partitioning Analysis for Probability of Bankruptcy

Panel A: Predictive ability

	(1)	(2)	(3)	(4)	(5)
Variables	$ln\left(\frac{V}{X}\right), Exret, \ln(E), \sigma_{AI}^{\omega}$	$ln\left(\frac{V}{X}\right), Exret, \ln(E), \sigma_{AI}^{\omega}, \sigma_F$	$ln\left(\frac{V}{X}\right), Exret, \ln(E), \sigma_{AI}^{\omega}, P95P5$	$ln\left(\frac{V}{X}\right), Exret, \ln(E), \sigma_{AI}^{\omega}, \sigma_F^{NAIVE}$	$ln\left(\frac{V}{X}\right), Exret, \ln(E), \sigma_{AI}^{\omega}, \sigma_F^{NAIVE(2)}$
AUC (Learning sample)	0.9299	0.9515	0.9417	0.9678	0.9587
AUC (Test sample)	0.9057	0.9165	0.9181	0.9258	0.9229
P5 (AUC(k)-AUC(1))		0.0000	-0.0019	0.0107	0.0058
Relative cost	0.2670	0.2110	0.2180	0.1760	0.2160
Hosmer-Lemeshow Test					
Test-statistic/ p-value	23.25/ 0.003	28.13/ <0.001	29.90/ <0.001	50.76/ <0.001	38.69/ <0.001

Panel B: Variable Importance

	(1)		(2)		(3)		(4)		(5)	
	Total	Primary Splitters	Total	Primary Splitters	Total	Primary Splitters	Total	Primary Splitters	Total	Primary Splitters
$ln\left(\frac{V}{X}\right)$	100	100	100	100	100	100	100	100	100	100
$Exret$	85.16	16.88	90.13	24.88	89.62	19.25	87.12	13.27	86.96	12.52
$\ln(E)$	39.50	14.22	37.03	0.69	39.40		39.91	0.81	38.66	1.98
σ_{AI}^{ω}	32.39	1.97	45.41	10.97	35.98	15.61	39.91	12.86	37.71	13.39
σ_F			59.06	31.76						
$P95P5$					50.86	21.03				
σ_F^{NAIVE}							75.41	37.85		
$\sigma_F^{NAIVE(2)}$									69.05	40.99

We use binary recursive partitioning, i.e. the Classification and Regression Trees methodology (CART) (Breiman, Friedman, Olshen and Stone, 1984) to create a decision tree that classifies firm-years into bankrupt or non-bankrupt. We follow the GINI rule to choose the optimal split at each node of the tree. Based on this approach, we generate the maximal tree and a set of sub-trees. We then use 10-fold cross validation to estimate the area under the ROC curve (AUC) for the different sub-trees and retain the tree that maximizes the AUC. Panel A reports summary statistics for the predictive ability of the model. P5 (AUC(k)-AUC(1)) is the 5th percentile of the difference between the AUC of each augmented model and the AUC of the base model. We use bootstrap resampling to calculate this statistic. Relative cost is the sum of the percentage of type I and type II errors. The Hosmer-Lemeshow test statistic is calculated as $\sum_{g=1}^{G} \frac{(O_g - E_g)^2}{N_g \pi_g (1-\pi_g)}$, where O_g, E_g, N_g and π_g are observed events, expected events, observations and predicted risk for group g and G is the number of groups. Panel B presents the importance scores for the variables in the model. These scores are calculated as the sum of the improvement that can be attributed to a given variable at each node of the tree. Total variable importance takes into account the role of the variable as a surrogate, while the column primary splitter only takes into account the role of the variable as a primary splitter. The analysis is based on a sample of 68,104 firm-months for the period January 1996 through to December 2012.

Table 4

Pooled regression of credit spreads on components of theoretical spreads: unconstrained analysis

$$OAS_{it} = \alpha_1 ln\left(\frac{V_{it}}{X_{it}}\right) + \alpha_2 Exret_{it} + \alpha_3 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+3}\sigma_{k,it} + \Gamma Control_{it} + \varepsilon_{it} \quad (8)$$

Panel A: Regression analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$ln\left(\frac{V}{X}\right)$	-0.004***	-0.004***	-0.005***	-0.005***	-0.004***	-0.003***	-0.000	-0.002***	-0.003***	-0.003***	-0.002***	-0.002***
	(-8.59)	(-8.76)	(-7.12)	(-7.04)	(-5.60)	(-4.80)	(-0.77)	(-4.87)	(-5.91)	(-5.89)	(-5.24)	(-5.13)
$Exret$	-0.010***	0.003	-0.005	-0.005	-0.005	-0.004	-0.013***	-0.006**	-0.005**	-0.005**	-0.006**	-0.005*
	(-2.69)	(1.19)	(-0.85)	(-0.85)	(-0.92)	(-0.72)	(-3.94)	(-2.57)	(-2.46)	(-2.47)	(-2.53)	(-1.96)
$ln(E)$	-0.002***	-0.002***	-0.003***	-0.003***	-0.003***	-0.003***	-0.004***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(-5.94)	(-4.60)	(-6.37)	(-6.33)	(-6.03)	(-5.43)	(-9.50)	(-8.11)	(-8.37)	(-8.35)	(-8.17)	(-7.25)
$Rating$	0.002***	0.002***	0.004***	0.004***	0.004***	0.004***	0.002***	0.001***	0.001***	0.001***	0.001***	0.001***
	(9.62)	(9.05)	(18.93)	(18.83)	(17.53)	(15.95)	(13.34)	(6.33)	(6.04)	(6.07)	(5.85)	(6.46)
Age	0.000***	0.000***	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**	0.000**
	(2.61)	(3.05)	(2.55)	(2.54)	(2.57)	(2.47)	(2.08)	(2.50)	(2.53)	(2.53)	(2.52)	(2.38)
$Duration$	0.000	0.000	0.000*	0.000*	0.000*	0.000	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(1.41)	(1.05)	(1.71)	(1.73)	(1.71)	(1.12)	(-6.44)	(-5.22)	(-5.16)	(-5.16)	(-5.22)	(-4.49)
σ_E	0.073***											
	(13.17)											
σ_I		0.093***						0.061***	0.059***	0.060***	0.060***	0.060***
		(14.91)						(13.06)	(13.11)	(13.10)	(13.04)	(11.86)
σ_F			0.106***						0.030***			
			(6.37)						(3.70)			
$P95P5$				0.030***						0.008***		
				(6.24)						(3.62)		
σ_F^{NAIVE}					0.024***						0.007	
					(2.86)						(1.25)	
$\sigma_F^{NAIVE(2)}$						0.017***						0.002
						(4.87)						(1.11)
σ_D							0.182***	0.135***	0.133***	0.133***	0.135***	0.122***
							(9.97)	(8.19)	(8.12)	(8.13)	(8.23)	(7.24)

Panel A (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	67,848	67,848	67,848	67,848	67,848	60,300	67,848	67,848	67,848	67,848	67,848	60,300
R2	0.641	0.671	0.568	0.567	0.561	0.576	0.692	0.732	0.733	0.733	0.732	0.729

Panel B: Marginal effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$ln\left(\frac{V}{X}\right)$	-0.1088	-0.1036	-0.1333	-0.1311	-0.0931	-0.0752	-0.0081	-0.0471	-0.0658	-0.0653	-0.0545	-0.0521
$Exret$	-0.0389	0.0129	-0.0182	-0.0184	-0.0203	-0.0169	-0.0537	-0.0225	-0.0214	-0.0215	-0.0221	-0.0192
$ln(E)$	-0.1290	-0.0914	-0.1660	-0.1654	-0.1628	-0.1575	-0.1935	-0.1416	-0.1449	-0.1447	-0.1442	-0.1344
$Rating$	0.2313	0.2042	0.4733	0.4779	0.4845	0.4773	0.2735	0.1349	0.1290	0.1297	0.1290	0.1489
Age	0.0268	0.0300	0.0299	0.0299	0.0304	0.0291	0.0188	0.0219	0.0221	0.0221	0.0221	0.0210
$Duration$	0.0157	0.0116	0.0205	0.0207	0.0206	0.0136	-0.1042	-0.0777	-0.0767	-0.0767	-0.0776	-0.0715
σ_E	0.5388											
σ_I		0.6181						0.4018	0.3939	0.3947	0.3955	0.3989
σ_F			0.1424						0.0406			
$P95P5$				0.1358						0.0390		
σ_F^{NAIVE}					0.0927						0.0284	
$\sigma_F^{NAIVE(2)}$						0.0527						0.0076
σ_D							0.5652	0.4174	0.4128	0.4132	0.4170	0.3779

Variable definitions are provided in Appendix I. Standard errors are clustered by firm and month. Regressions are based on a sample of 67,848 firm-months for the period January 1996 through to December 2012.

Marginal effects are reported as the marginal increase in option adjusted credit spreads as each of the explanatory variables increases by one standard deviation.

Table 5
Cross-sectional partitions: unconstrained analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
HY	0.030***	0.002	0.081***	0.082***	0.085***	0.087***	0.060***	0.012**	0.013***	0.013***	0.014***	0.014**
	(4.25)	(0.25)	(9.02)	(9.10)	(9.48)	(8.32)	(9.98)	(2.33)	(2.60)	(2.60)	(2.81)	(2.29)
$ln\left(\frac{V}{X}\right) * HY$	-0.009***	-0.006***	-0.016***	-0.016***	-0.016***	-0.015***	-0.004***	-0.002**	-0.003***	-0.003***	-0.004***	-0.004***
	(-7.18)	(-5.98)	(-8.92)	(-8.87)	(-8.64)	(-8.08)	(-3.72)	(-2.55)	(-3.18)	(-3.17)	(-3.65)	(-3.88)
$Exret * HY$	-0.013***	-0.001	0.002	0.002	0.002	0.002	-0.013***	-0.010***	-0.010***	-0.010***	-0.010***	-0.009***
	(-3.37)	(-0.20)	(0.38)	(0.39)	(0.36)	(0.45)	(-4.10)	(-3.57)	(-3.55)	(-3.55)	(-3.56)	(-2.94)
$ln(E) * HY$	-0.003***	-0.002**	-0.007***	-0.007***	-0.007***	-0.007***	-0.006***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(-3.57)	(-2.33)	(-7.03)	(-7.06)	(-7.16)	(-6.24)	(-8.58)	(-4.15)	(-4.26)	(-4.26)	(-4.21)	(-3.44)
$Age * HY$	0.001*	0.000*	0.001**	0.001**	0.001**	0.001***	0.001**	0.000*	0.000*	0.000*	0.000*	0.001*
	(1.90)	(1.69)	(2.22)	(2.22)	(2.58)	(2.75)	(2.20)	(1.68)	(1.68)	(1.68)	(1.80)	(1.93)
$Duration * HY$	-0.000	-0.001	-0.000	-0.000	-0.000	-0.000	-0.001**	-0.001*	-0.001*	-0.001*	-0.001*	-0.001
	(-1.06)	(-1.50)	(-0.21)	(-0.18)	(-0.20)	(-0.70)	(-2.10)	(-1.74)	(-1.70)	(-1.70)	(-1.73)	(-1.61)
$\sigma_E * HY$	0.047***											
	(7.61)											
$\sigma_I * HY$		0.069***						0.047***	0.045***	0.045***	0.044***	0.049***
		(11.46)						(8.07)	(7.83)	(7.84)	(7.79)	(7.63)
$\sigma_F * HY$			0.170***						0.035**			
			(6.69)						(2.50)			
$P95P5 * HY$				0.048***						0.010**		
				(6.60)						(2.45)		
$\sigma_F^{NAIVE} * HY$					0.052***						0.020*	
					(3.95)						(1.96)	
$\sigma_F^{NAIVE(2)} * HY$						0.047***						0.012**
						(5.36)						(1.98)
$\sigma_D * HY$							0.045*	-0.000	-0.002	-0.002	-0.000	-0.015
							(1.66)	(-0.02)	(-0.06)	(-0.06)	(-0.01)	(-0.53)
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	67,848	67,848	67,848	67,848	67,848	60,300	67,848	67,848	67,848	67,848	67,848	60,300
R2	0.668	0.708	0.620	0.619	0.617	0.632	0.712	0.751	0.752	0.752	0.753	0.752

Variable definitions are provided in Appendix I. Standard errors are clustered by firm and month. The full regression includes all the main effects in addition to the reported interaction terms. Regressions are based on a sample of 67,848 firm-months for the period January 1996 through to December 2012.

Table 6
Pooled regression of credit spreads on theoretical credit spreads: constrained analysis

$$OAS_{it} = \alpha_1 Exret_{it} + \alpha_2 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+2}\, CS_{\sigma_{k,it}} + \Gamma Control_{it} + \varepsilon_{it} \quad (9)$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$Exret$	-0.010***	-0.007***	-0.003	-0.009***	-0.006***	-0.002	-0.004**	-0.002	-0.004**	-0.003	-0.004**	-0.003	-0.005**
	(-3.58)	(-3.36)	(-1.18)	(-4.84)	(-2.96)	(-0.70)	(-2.20)	(-0.68)	(-2.19)	(-1.06)	(-2.40)	(-0.96)	(-2.38)
$ln(E)$	-0.002***	-0.002***	-0.002***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(-6.15)	(-5.88)	(-5.43)	(-9.10)	(-8.82)	(-7.92)	(-8.96)	(-7.74)	(-8.92)	(-6.30)	(-8.41)	(-5.32)	(-7.82)
$Rating$	0.003***	0.003***	0.003***	0.002***	0.002***	0.003***	0.002***	0.003***	0.002***	0.003***	0.002***	0.003***	0.002***
	(16.04)	(15.72)	(16.48)	(13.54)	(15.15)	(16.73)	(14.94)	(16.79)	(15.03)	(16.41)	(15.31)	(14.58)	(13.89)
Age	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000*	0.000
	(1.48)	(0.81)	(0.53)	(0.59)	(0.24)	(0.67)	(0.04)	(0.65)	(0.03)	(1.30)	(0.31)	(1.73)	(0.56)
$Duration$	-0.000	-0.000	-0.000	-0.000	-0.000*	-0.000	-0.000*	-0.000	-0.000*	-0.000	-0.000**	-0.000	-0.000*
	(-0.02)	(-0.72)	(-1.22)	(-1.49)	(-1.83)	(-0.58)	(-1.84)	(-0.50)	(-1.82)	(-1.13)	(-2.04)	(-1.24)	(-1.90)
$CS_{\sigma_E}^{BASE}$	0.601***												
	(12.90)												
CS_{σ_E}		0.663***											
		(15.86)											
CS_{σ_I}			0.797***										
			(16.30)										
$CS_{\sigma_A^{\omega}}$				0.675***									
				(16.92)									
$CS_{\sigma_{AI}^{\omega}}$					0.796***		0.578***		0.584***		0.606***		0.636***
					(17.98)		(14.19)		(14.23)		(14.52)		(12.72)
CS_{σ_F}						0.703***	0.275***						
						(15.67)	(8.70)						
CS_{P95P5}								0.702***	0.270***				
								(15.34)	(8.34)				
$CS_{\sigma_F^{NAIVE}}$										0.758***	0.266***		
										(13.41)	(5.57)		
$CS_{\sigma_F^{NAIVE}}(2)$												0.955***	0.249***
												(10.81)	(4.39)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 6 (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Nobs	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	50,407	50,407
R2	0.706	0.757	0.756	0.773	0.774	0.736	0.786	0.734	0.786	0.727	0.784	0.704	0.773
R2(U)	0.633	0.633	0.671	0.591	0.603	0.559	0.611	0.558	0.610	0.552	0.605	0.566	0.609
Included σ	σ_E	σ_E	σ_I	σ_A^{ω}	σ_{AI}^{ω}	σ_F	$\sigma_{AI}^{\omega}, \sigma_F$	$P5P95$	$\sigma_{AI}^{\omega}, P95P5$	σ_F^{NAIVE}	$\sigma_{AI}^{\omega}, \sigma_F^{NAIVE}$	$\sigma_F^{NAIVE(2)}$	$\sigma_{AI}^{\omega}, \sigma_F^{NAIVE(2)}$
Vuong Z	20.903	22.852	19.013	29.548	25.251	24.616	24.981	24.400	24.990	23.033	24.687	19.401	20.901
p-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Variable definitions are provided in Appendix I and calculation of theoretical credit spreads is described in Appendix III. Standard errors are clustered by firm and month. R2(U) is the R-square from the estimation of equation (8) with the volatility measures included in each of the constrained specifications (Included σ). Regressions are based on a sample of 56,846 firm-months for the period June 1999 through to December 2012 (we lose the part of the sample to ensure that our bankruptcy forecasts are 'out of sample').

Table 7
Cross-sectional partitions: constrained analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
HY	0.037***	0.029***	0.033***	0.039***	0.045***	0.052***	0.044***	0.053***	0.044***	0.052***	0.044***	0.051***	0.046***
	(6.43)	(5.64)	(6.03)	(8.73)	(9.42)	(8.90)	(9.24)	(8.91)	(9.27)	(8.82)	(9.11)	(7.40)	(8.26)
$Exret * HY$	-0.011***	-0.009***	-0.006**	-0.012***	-0.011***	-0.003	-0.009***	-0.003	-0.009***	-0.004	-0.010***	-0.004	-0.010***
	(-3.85)	(-3.54)	(-2.33)	(-4.95)	(-4.65)	(-1.00)	(-4.06)	(-0.98)	(-4.07)	(-1.24)	(-4.14)	(-1.03)	(-3.96)
$ln(E) * HY$	-0.003***	-0.002***	-0.002***	-0.003***	-0.004***	-0.005***	-0.004***	-0.005***	-0.004***	-0.005***	-0.004***	-0.004***	-0.004***
	(-4.35)	(-2.92)	(-3.37)	(-6.17)	(-6.86)	(-6.62)	(-6.60)	(-6.64)	(-6.62)	(-6.45)	(-6.46)	(-5.42)	(-5.88)
$Age * HY$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001*	0.000
	(1.25)	(0.48)	(0.35)	(0.43)	(0.29)	(0.45)	(0.12)	(0.45)	(0.12)	(1.00)	(0.34)	(1.69)	(0.79)
$Duration *$	-0.001*	-0.001**	-0.001***	-0.001**	-0.001***	-0.001***	-0.001***	-0.001**	-0.001***	-0.001**	-0.001***	-0.001***	-0.001***
HY	(-1.65)	(-2.56)	(-3.00)	(-2.47)	(-2.59)	(-2.58)	(-2.99)	(-2.47)	(-2.95)	(-2.48)	(-2.90)	(-3.20)	(-3.04)
$CS_{\sigma_E}^{BASE} * HY$	0.278***												
	(5.46)												
$CS_{\sigma_E} * HY$		0.288***											
		(5.76)											
$CS_{\sigma_i} * HY$			0.333***										
			(5.70)										
$CS_{\sigma_A^{\omega}} * HY$				0.212***									
				(4.31)									
$CS_{\sigma_{AI}^{\omega}} * HY$					0.228***		0.022		0.028		0.048		0.098
					(3.85)		(0.37)		(0.46)		(0.67)		(1.23)
$CS_{\sigma_F} * HY$						0.395***	0.195***						
						(7.30)	(3.78)						
$CS_{P95P5} * HY$								0.399***	0.190***				
								(7.26)	(3.63)				
$CS_{\sigma_F^{NAIVE}} * HY$										0.451***	0.200***		
										(7.41)	(2.92)		
$CS_{\sigma_F^{NAIVE(2)}} *$												0.635***	0.177*
HY												(7.61)	(1.91)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	56,846	50,407	50,407
R2	0.700	0.752	0.750	0.774	0.775	0.742	0.787	0.740	0.786	0.736	0.785	0.720	0.775

Variable definitions are provided in Appendix I and calculation of theoretical credit spreads is described in Appendix III. Standard errors are clustered by firm and month. The full regression includes all the main effects in addition to the reported interaction terms. Regressions are based on a sample of 56,846 firm-months for the period June 1999 through to December 2012 (we lose the first part of the sample to ensure that our bankruptcy forecasts are 'out of sample').

Table 8

Robustness : 5 year CDS spread analysis

$$CDS5Y_{it} = \alpha_1 Exret_{it} + \alpha_2 ln(E_{it}) + \sum_{k=1}^{K} \alpha_{k+2}\, CS_{\sigma_{k,it}} + \Gamma Control_{it} + \varepsilon_{it} \quad (10)$$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
$Exret$	-0.005	-0.005	-0.002	-0.010*	-0.007	0.002	-0.004	0.002	-0.004	0.005	-0.003	0.009	0.000
	(-0.51)	(-0.60)	(-0.29)	(-1.65)	(-1.09)	(0.29)	(-0.68)	(0.30)	(-0.67)	(0.61)	(-0.63)	(1.13)	(0.05)
$ln(E)$	-0.001	0.000	0.000	-0.000	-0.000	-0.001	-0.000	-0.001	-0.000	-0.001	-0.000	-0.002	-0.001
	(-0.58)	(0.22)	(0.37)	(-0.41)	(-0.45)	(-0.55)	(-0.05)	(-0.53)	(-0.04)	(-0.81)	(-0.18)	(-1.59)	(-0.93)
$Rating$	0.004***	0.003***	0.003***	0.003***	0.003***	0.004***	0.003***	0.004***	0.003***	0.004***	0.003***	0.003***	0.003***
	(6.47)	(7.36)	(7.73)	(8.56)	(8.48)	(9.26)	(8.49)	(9.30)	(8.53)	(8.40)	(8.18)	(8.60)	(8.59)
$CS_{\sigma_E}^{BASE}$	0.965***												
	(6.59)												
CS_{σ_E}		1.189***											
		(7.57)											
CS_{σ_I}			1.106***										
			(7.46)										
$CS_{\sigma_A^{\omega}}$				1.015***									
				(7.13)									
$CS_{\sigma_{AI}^{\omega}}$					0.910***		0.693***		0.689***		0.720***		0.741***
					(7.13)		(5.56)		(5.67)		(5.21)		(4.97)
CS_{σ_F}						1.020***	0.404***						
						(6.85)	(4.42)						
CS_{P95P5}								1.007***	0.404***				
								(6.75)	(4.46)				
$CS_{\sigma_F^{NAIVE}}$										1.003***	0.367***		
										(9.09)	(2.76)		
$CS_{\sigma_F^{NAIVE}}(2)$												1.135***	0.243**
												(5.54)	(2.08)
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Nobs	22,199	22,199	22,199	22,199	22,199	22,199	22,199	22,199	22,199	22,199	22,199	18,899	18,899

Table 8 (Cont.)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
R2	0.385	0.438	0.462	0.522	0.523	0.477	0.535	0.479	0.536	0.468	0.533	0.453	0.541
R2(U)	0.406	0.406	0.463	0.350	0.357	0.355	0.373	0.354	0.371	0.339	0.358	0.353	0.362
Included σ	σ_E	σ_E	σ_I	σ_A^{ω}	σ_{AI}^{ω}	σ_F	$\sigma_{AI}^{\omega}, \sigma_F$	$P5P95$	$\sigma_{AI}^{\omega}, P95P5$	σ_F^{NAIVE}	$\sigma_{AI}^{\omega}, \sigma_F^{NAIVE}$	$\sigma_F^{NAIVE(2)}$	$\sigma_{AI}^{\omega}, \sigma_F^{NAIVE(2)}$
Vuong Z	-5.635	4.577	-0.233	8.013	7.996	7.748	7.745	7.679	7.741	6.346	7.533	5.379	6.206
p-value	<0.0001	<0.0001	0.8153	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

Variable definitions are provided in Appendix I and calculation of theoretical credit spreads is described in Appendix III. Standard errors are clustered by firm and month. Regressions are based on a sample of 22,199 firm-months for the period January 2004 through to December 2012 where we have available CDS data.