

What Drives the Cross-Section of Credit Spreads?: A Variance Decomposition Approach*

Yoshio Nozawa[†]

February 9, 2016

Abstract

I decompose the variation of credit spreads for corporate bonds into changing expected returns and changing expectation of credit losses. Using a log-linearized pricing identity and a vector autoregression applied to micro-level data from 1973 to 2011, I find that expected returns contribute to the cross-sectional variance of credit spreads nearly as much as expected credit loss does. However, most of the time-series variation in credit spreads for the market portfolio corresponds to risk premiums. The joint decomposition of bonds and stocks show that expected credit loss for bonds are negatively correlated with expected returns and cash flows for stocks. At the firm level, both risk premiums and credit loss negatively affect corporate investments.

*I am grateful for comments and suggestions from John Cochrane, Joost Driessen, Simon Gilchrist, Lars Hansen, Don Kim, Ralph Koijen, Arvind Krishnamurthy, Marcelo Ochoa, Pietro Veronesi, Bin Wei, Kenji Wada, Vladimir Yankov and participants in workshops at Chicago Booth, Chicago Fed, Erasmus Credit Conference, FRB and New York Fed. The views expressed herein are the author's and do not necessarily reflect those of the Board of Governors of the Federal Reserve System.

[†]Division of Monetary Affairs, Board of Governors of the Federal Reserve System, Mail Stop 165, 20th Street and Constitution Avenue NW, Washington, DC 20551. Email: yoshio.nozawa@frb.gov.

1 Introduction

What drives the cross-sectional variation in credit spreads? Credit spreads are higher when the issuer of a corporate bond faces a higher risk of default and when the rate at which the corporate bond's cash flows are discounted rises. Since the expected default and expected returns are unobservable, past research often relies on structural models of debt, such as the Merton (1974) model, to decompose credit spreads. However, there is little agreement on the best measures of expected default loss and expected returns. In this article, I take advantage of a large panel dataset of the US corporate bond prices and estimate the conditional expectations without relying on a particular model of default. Based on these estimates, I quantify the contributions of the default component and the discount rate component to the credit spread variation.

I apply the variance decomposition approach of Campbell and Shiller (1988a and 1988b) to the credit spread. In the decomposition, the credit spread plays the role of the dividend-price ratio for stocks, while credit loss plays the role of dividend growth. This decomposition framework relates the current credit spread to the sum of expected excess returns and credit losses over the long run. This relationship implies that, if the credit spread varies, then either long-run expected excess returns or long-run expected credit loss must vary.

I estimate a VAR involving credit spreads, excess returns, probability of default and credit rating of the corporate bonds. Since default occurs infrequently, estimating the expected credit loss and expected returns by running forecasting regressions requires a large dataset. Therefore, I collect corporate bond prices from the Lehman Brothers Fixed Income Database, the Mergent FISD/NAIC Database, TRACE and DataStream, which provide an extensive dataset of the publicly traded corporate bonds from 1973 to 2011. In addition, I use Moody's Default Risk Service to make sure that the price observations upon default are complete, and thus my credit loss measure does not miss bond defaults that occur during the sample period.

Based on the estimated VAR, I find that the ratio of volatility of the implied long-run expected credit loss to the volatility of credit spreads is 0.67, while the ratio of the risk premium volatility to the credit spread volatility is 0.52. In the world where the credit spreads are driven solely by the expected default, the volatility ratio for the expected credit loss would be one, while the ratio for the risk premium would be zero. Instead, about half of the volatility of credit spreads comes from changing expected excess returns. In contrast, I show that the drivers for the market-wide variation in credit spreads are mostly risk premiums, unlike those for the individual bonds. The difference arises due to the diversification effects: The default shocks are more idiosyncratic than the expected return shocks, and thus the expected credit loss component is more important at the individual bond level than at the aggregate market level.

I find a non-linear relationship among risk premiums, expected credit loss and credit spreads, depending on the credit rating of the bond. Much of the variation in credit spreads within investment grade (IG) bonds corresponds to the risk premium variation, while the expected credit loss accounts for a larger fraction of the volatility of the credit spreads of high yield bonds.

Next, I extend the variance decomposition framework to study the interaction in expected cash flows and risk premiums between bonds and stocks. To study the interaction, I jointly decompose the cross-section of bond and stock prices, and find a significant positive correlation in expected cash flows between bonds and stocks, while the risk premium correlation is insignificant. Interestingly, the correlation between the expected default on bonds and the risk premium on stocks is negative. Thus, my VAR specification yields the results consistent with the distress anomaly of Campbell, Hilscher and Szilagyi (2008), who find that a stock of a firm near default earns lower expected returns.

The previous literature shows the strong link between credit spreads and economic activity (e.g. Philippon (2009), Gilchrist, Yankov and Zakrajšek (2009), and Gilchrist, Sim

and Zakrajšek (2013)). Motivated by their findings, I forecast corporate investment using the two components of credit spreads identified in the previous analysis. Using the firm-level panel data, I regress the ratio of investment to capital this year on the expected credit loss and excess returns at the end of the previous year. At the individual firm level, both components negatively forecast the investment rates even after controlling for the traditional proxies for Tobin's q . These results are in stark contrast with the results at the aggregate level, in which the risk premium variation dominates the expected credit loss variation in forecasting investments.

Related Literature

The papers closest to mine are Bongaerts (2010) and Elton, Gruber, Agrawal and Mann (2001). The idea of applying a variance decomposition approach to corporate bonds starts in Bongaerts (2010), who decomposes variance of the returns on the corporate bond indices. This article is a complement to Bongaerts (2010), as I use micro-level data to study the cross-section of corporate bonds, and decompose credit spreads rather than returns. Elton, Gruber, Agrawal and Mann (2001) explain the level of the average credit spreads for AA, A and BBB bonds based on the average probability of default and loss given default. In contrast, this article decomposes the variance of credit spreads allowing for the time-varying probability of default and risk premiums. By studying the variance of the credit spreads, I show a link in movements between the different components of the credit spreads and the issuers' investment.

In addition, numerous papers explain the credit spread using structural models of debt (e.g., Leland (1994), Collin-Dufresne and Goldstein (2001), Collin-Dufresne, Goldstein and Martin (2001), Chen, Collin-Dufresne and Goldstein (2009), Bharmra, Kuehn and Strebulaev (2010), Chen (2010), and Huang and Huang (2012)), reduced-form models (Duffee (1999) and Driessen (2005)) or the credit default swap spreads (Longstaff, Mithal and Neis (2005)). This article differs from the literature as I do not make assumptions about how firms make

their decisions about their capital structure and defaults, or on what factors drive the firm value.

This article examines the contribution of variation in expected returns on corporate bond prices, which complements the excess volatility and return predictability found in stock prices (e.g., Campbell and Shiller, (1988a and 1998b), Campbell, (1991), Vuolteenaho, (2002), and Cochrane, (2008 and 2011)). In addition, this paper adds to the literature which studies the information content in the price ratios of a variety of assets. For Treasury bonds, Fama and Bliss (1988) and Cochrane and Piazzesi (2005) find that forward rates forecast bond returns, not future short rates. For foreign exchange, Hansen and Hodrick (1980), Fama (1984), and Lustig and Verdelhan (2007) show that uncovered interest rate parity does not hold in the data. Beber (2006), McAndrews (2008), Taylor (2009) and Schwartz (2013) decompose the yield spreads in the sovereign and money markets.

The rest of the article is organized as follows: Section 2 shows the decomposition of the credit spread of corporate bonds. I describe the data and show the empirical results in Section 3. Section 4 presents a joint variance decomposition of bonds and stocks. Section 5 examines the variance decomposition at the market level, and the effect of risk premiums and expected credit losses on firms' investment decisions. Section 6 provides concluding remarks.

2 Decomposition of Corporate Bond Credit Spreads

2.1 Log-linear Approximation of Bond Excess Returns

I log-linearize excess returns on a corporate bond to obtain a linear relationship among log excess returns, credit spreads and credit loss. I consider the strategy where an investor buys and holds an individual corporate bond i until it matures or defaults. If the bond defaults, the investor sells the defaulted bond and buys the Treasury bond with the same coupon rate

and remaining time to maturity as the defaulting bond.

Let $P_{i,t}$ be the price per one dollar face value for corporate bond i at time t including accrued interest, and $C_{i,t}$ be the coupon rate. Then, the return on the bond is

$$R_{i,t+1} = \frac{P_{i,t+1} + C_{i,t+1}}{P_{i,t}}.$$

Suppose that there is a matching Treasury bond for corporate bond i , such that the matching Treasury bond has an identical coupon rate and repayment schedule as corporate bond i . Let $P_{i,t}^f$ and $C_{i,t}^f$ be the price and coupon rate for such a Treasury bond. Then, the return on the matching Treasury bond is

$$R_{i,t+1}^f = \frac{P_{i,t+1}^f + C_{i,t+1}^f}{P_{i,t}^f}.$$

As I do not have the data for the loss upon default for coupon payments, I assume that the rate of credit loss (defined below) for the coupons is the same as the rate for the principal. I log-linearize both $R_{i,t+1}$ and $R_{i,t+1}^f$ using the same expansion point, $\rho \in [0, 1)$.

The log return on corporate bond i , in excess of the log return on the matching Treasury bond, can then be approximated as

$$r_{i,t+1}^e \equiv \log R_{i,t+1} - \log R_{i,t+1}^f \approx -\rho s_{i,t+1} + s_{i,t} - l_{i,t+1} + \text{const}, \quad (1)$$

where

$$s_{i,t} \equiv \begin{cases} \log \frac{P_{i,t}^f}{P_{i,t}} & \text{if } t < t_D, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

$$l_{i,t} \equiv \begin{cases} \log \frac{P_{i,t}^f}{P_{i,t}} & \text{if } t = t_D, \\ 0 & \text{otherwise,} \end{cases} \quad (3)$$

where t_D is the time of default. The variable $s_{i,t}$ measures credit spreads while $l_{i,t}$ measures the credit loss upon default. Equation (1) implies that the excess return on corporate bond i is low due to either widening credit spreads or defaults. In Appendix A, I show the detailed derivation of (1).

The credit spread measure, $s_{i,t}$, is the price spreads rather than a yield spread. Price spreads have important advantages over yield spreads: The price spread has a definition based on a simple formula. Therefore, $s_{i,t}$ can be approximated using a linear function of $s_{i,t+1}$, $r_{i,t+1}^e$ and $l_{i,t+1}$ in (1) without inducing large approximation errors. In contrast, yield spreads for coupon bearing bonds can only be defined implicitly and computed numerically, which makes it hard to express the bond returns using a linear function of yield spreads. However, the price spread, $s_{i,t}$, is closely related to the commonly used yield spread, since a price change can be approximated by a change in yields multiplied by duration.¹ Thus, both spreads are, conceptually and empirically, closely tied together, and the analysis on the price spreads is useful in understanding the information content in the yield spreads.

The credit loss measure, $l_{i,t}$, encodes the information about both the incidence of default and the loss given default. The loss given default is measured using the market price of the corporate bonds upon default. As such, this measure of loss given default is the loss for an investor who invests in corporate bonds. This measure of credit loss is consistent with the way in which Moody's estimates the loss given default,² which is widely used in pricing credit derivatives.

To determine if a bond is in default, I follow Moody's (2011) definition of defaults. A bond is in default if there is (a) missed or delayed repayments, (b) a bankruptcy filing or legal

¹The average cross-sectional correlation between the price spreads and the yield spreads in my sample is 0.82, while the correlation between the price spreads and the yield spreads times duration is 0.97.

²For example, Moody's (1999) reports "One methodology for calculating recovery rates would track all payments made on a defaulted debt instrument, discount them back to the date of default, and present them as a percentage of the par value of the security. However, this methodology, while not infeasible, presents a number of calculation problems and relies on a variety of assumptions.... For these reasons, we use the trading price of the defaulted instrument as a proxy for the present value of the ultimate recovery."

receivership that will likely cause a miss or delay in repayments, (c) a distressed exchange or (d) a change in payment terms that results in a diminished financial obligations for the borrower. The definition does not include so-called technical defaults, such as temporary violations of the covenants regarding financial ratios, and slightly delayed payments due to technical or administrative errors.

None of the variables on the right-hand side of (1) depend on the coupon payments. Since the corporate bond and the Treasury bond have the same coupon rates, the coupons cancel each other. As a result, there is no seasonality in these variables, enabling one to use monthly returns for the decomposition. Moreover, the decomposition results are not subject to the assumption of the cash flow reinvestment, a point emphasized in Chen (2009).³

The difference equation (1) approximates log excess returns using the first-order Taylor expansion. In the empirical work below, I set the value of ρ to be 0.992, which minimizes the approximation error in (1). I show below that the approximation error is small and does not affect my empirical results.

Now I iterate the difference equation forward up to the maturity of bond i , T_i . That is,

$$s_{i,t} \approx \sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e + \sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} + \text{const.} \quad (4)$$

If the bond defaults at $t_D < T_i$, the investor adjusts the position such that $r_{i,t}^e = l_{i,t} = 0$ for $t > t_D$. Therefore, I can still iterate the difference equation forward up to T_i with no consequences.

Since (4) holds path-by-path, the approximate equality holds under expectation. Taking

³Chen (2009) points out that the use of annual horizon makes it necessary to make an assumption about how the cash flows paid out in the middle of a year are reinvested by investors, and the variance decomposition results are sensitive to such assumptions. Since the coupon payments from the corporate bond and the Treasury bond offset with each other, the variance decomposition in this article does not rely on the assumption about cash flow reinvestments.

the time t conditional expectation of the both sides of (4), we have

$$s_{i,t} \approx E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e \middle| \mathcal{F}_t \right] + E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} \middle| \mathcal{F}_t \right] + const, \quad (5)$$

where \mathcal{F}_t is the information set of economic agents.

Equation (5) shows that the variation in credit spreads can be decomposed into long-run expected excess returns or credit loss without leaving unexplained residuals. The movement in credit spreads must forecast either excess returns or defaults. The basic idea behind this decomposition is the same as that behind the decomposition of the price-dividend ratio for a stock. Since corporate bonds have fixed cash flows, the only source of shocks to cash flows is credit loss. Thus, the term $l_{i,t}$ plays a role analogous to dividend growth for equities. In the case of corporate bonds, however, we have $s_{i,T_i} = 0$ by construction. As a result, I do not have to impose the condition in which $\rho^j s_{i,t+j}$ tends to zero, as j goes to infinity.

Let us define the long-run expected credit loss as

$$s_{i,t}^l \equiv E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} \middle| \mathcal{F}_t \right].$$

We can then measure how much the volatility of $s_{i,t}$ corresponds to the volatility of the expected credit loss by the ratio $\sigma(s_{i,t}^l) / \sigma(s_{i,t})$. To evaluate the magnitude of $\sigma(s_{i,t}^l) / \sigma(s_{i,t})$, it is useful to set a benchmark case, in which all volatility in the credit spread is associated with the expected credit loss.

Definition. The expected credit loss hypothesis holds if a change in the credit spread only reflects the news about the expected credit loss. That is,

$$s_{i,t} = s_{i,t}^l + const,$$

holds.

Under the expected credit loss hypothesis, $\sigma(s_{i,t}^l) / \sigma(s_{i,t}) = 1$ holds. Therefore, using the hypothesis as a benchmark, we can ask how far from one the estimated volatility ratio is. The expected credit loss hypothesis also implies that the long-run expected excess returns, $E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e \middle| \mathcal{F}_t \right]$, are constant.

The expected credit loss hypothesis is the corporate bond counterpart of the expectation hypothesis for interest rates and of uncovered interest rate parity for foreign exchange rates. These hypotheses share the same basic idea that the current scaled price should reflect the future fundamentals in an unbiased way. If these hypotheses fail, either due to time-varying risk premiums or irrational expectations, then the excess returns are forecastable using the scaled price.

3 Empirical Results

3.1 Data

I construct the panel data of corporate bond prices from the Lehman Brothers Fixed Income Database, the Mergent FISD/NAIC Database, TRACE and DataStream. Appendix B provides a detailed description of these databases. When there are overlaps among the four databases, I prioritize in the following order: the Lehman Brothers Fixed Income Database, TRACE, Mergent FISD/NAIC and DataStream. I check whether the main result is robust to the change in orders in Appendix B. If the observation is missing in the databases above, I use Moody's Default Risk Service to complement the price upon default. CRSP and Compustat provide the stock prices and accounting information.

I remove bonds with floating rates and with option features other than callable bonds. Until the late 1980s, very few bonds were non callable, and thus removing callable bonds would significantly reduce the length of the sample period. Crabbe (1991) estimates that call options contribute nine basis points to the bond spread, on average, for investment grade

bonds. Therefore, the effect of call options does not seem large enough to significantly affect my results. To show the robustness of the results, I include fixed effects for callable bonds, repeat the main exercise in the online appendix,⁴ and show that callability does not drive the main results.

I apply three filters to remove the observations that are likely to be subject to erroneous recording. First, I remove the price observations that are higher than matching Treasury bond prices. Second, I drop the price observations below one cent per dollar. Third, I remove the return observations that show a large bounceback. Specifically, I compute the product of the adjacent return observations and remove both observations if the product is less than -0.04 .

In order to compute excess returns and credit spreads, I construct the prices of the synthetic Treasury bonds that match the corporate bonds using the Federal Reserve’s constant-maturity yields data. The methodology is detailed in Appendix B.

3.2 Estimation by a VAR

I estimate the conditional expectations in (5) and measure their volatilities, based on a VAR. To focus on the cross-sectional variation, I subtract the cross-sectional mean at time t from the state variables, and denote them with tilde. In the basic setup, I use a vector of state variables,

$$X_{i,t} = \begin{pmatrix} \tilde{r}_{i,t}^e & d_{i,t} \tilde{s}_{i,t} & \tau_{i,t} \tilde{z}_{i,t} \end{pmatrix}', \quad (6)$$

where $d_{i,t}$ is a vector of dummy variables for credit ratings defined by $d_{i,t} = \begin{pmatrix} 1 & d_{i,t}^{Baa} & d_{i,t}^{Ba} & d_{i,t}^{B-} \end{pmatrix}$, and $d_{i,t}^\theta$ is the dummy for rating θ , $\tau_{i,t}$ is the bond’s duration and $z_{i,t}$ is a vector of state variables other than $\tilde{r}_{i,t}^e$ and $\tilde{s}_{i,t}$.

⁴The online appendix can be found at: <https://sites.google.com/site/yoshiofinancialeconomics/home/research>

The dynamics of the state variables is given by

$$X_{i,t+1} = AX_{i,t} + W_{i,t+1}. \quad (7)$$

Matrix A is held constant both over time and across bonds. This VAR specification implies that ex-ante, a bond is expected to behave similarly to other bonds with the same values of the state variables. I also assume that $W_{i,t}$ is independent over time but can be correlated across bonds.

To address the concern about the assumption of constant coefficient A , I allow two interaction terms to better capture the dynamics.

First, by interacting $\tilde{s}_{i,t}$ with $d_{i,t}$, I allow the VAR coefficient to change based on the bond's credit rating. I show later that there is a significant non-linearity between expected credit loss and credit spreads, which is well captured by this interaction term.

Second, since many structural models of debt (e.g., Merton (1974)) or reduced form models (e.g., Duffie and Singleton (1999)) imply that the expected returns and the risk of a corporate bond depend on its time to maturity, the state variables $z_{i,t}$ are scaled by the bond's duration. The price spread, $\tilde{s}_{i,t}$, has a convenient feature in that it tends to shrink with its duration: Since a price spread is roughly equal to a yield spread times the bond's duration, holding yield spreads constant, $\tilde{s}_{i,t}$ tends to zero as the bond approaches maturity. Thus, I do not scale $\tilde{s}_{i,t}$ with duration.

Let $e_i, i = 1, 2$ be unit vectors whose i -th entry is one while the other entries are zero. Then, the long-run expected loss and excess returns implied by the VAR is

$$\tilde{s}_{i,t}^l = E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} \tilde{l}_{i,t+j} \middle| X_{i,t} \right] = e_L G(T_i) X_{i,t}, \quad (8)$$

$$\tilde{s}_{i,t}^r = E \left[\sum_{j=1}^{T_i-t} \rho^{j-1} \tilde{r}_{i,t+j}^e \middle| X_{i,t} \right] = e_1 G(T_i) X_{i,t}, \quad (9)$$

where $G(T_i) \equiv A(I - \rho A)^{-1} \left(I - (\rho A)^{T_i-t} \right)$ and $e_L = -\rho e_2 + e_2 A^{-1} - e_1$.⁵

Since we condition on $X_{i,t} \subseteq \mathcal{F}_{i,t}$, the estimated volatilities based on the VAR, $\sigma(\tilde{s}_{i,t}^l)$ and $\sigma(\tilde{s}_{i,t}^r)$, give the lower bound for the true volatility based on the agent's information set.

By identity (4),

$$e_1 G(T_i) + e_L G(T_i) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \end{pmatrix} \quad (10)$$

holds. Moreover, the expected credit loss hypothesis implies

$$\begin{aligned} e_1 G(T_i) &= \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \\ e_L G(T_i) &= \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \end{pmatrix}, \end{aligned}$$

must hold.

Unlike the forecasting coefficients in (10), the volatility ratios for expected credit loss, $\sigma(\tilde{s}_{i,t}^l) / \sigma(s_{i,t})$, and expected excess returns, $\sigma(\tilde{s}_{i,t}^r) / \sigma(s_{i,t})$, do not have to add up to one, due to the covariance between expected credit loss and expected excess returns.

For statistical inference, I compute the standard errors of the VAR-implied long-run coefficients and volatility ratios by the delta method. To this end, I numerically calculate the derivative of the long-run coefficients and volatility ratios with respect to the VAR parameters.

⁵To obtain (8), I use the one-period identity in (1). Solving for $\tilde{l}_{i,t+1}$ and taking the conditional expectation, we have

$$\begin{aligned} E \left[\tilde{l}_{i,t+j} \middle| X_{i,t} \right] &= E \left[-\rho e_2 X_{i,t+j} + e_2 X_{i,t+j-1} - e_1 X_{i,t+j} \middle| X_{i,t} \right], \\ &= e_L A^j X_{i,t}. \end{aligned}$$

Plugging $E \left[\tilde{l}_{i,t+j} \middle| X_{i,t} \right]$ into $E \left[\sum \rho^{j-1} \tilde{l}_{i,t+j} \middle| X_{i,t} \right]$ yields (8).

3.3 Main Results

In this section, I estimate the VAR in (7) and quantify the contribution of the volatility of expected credit loss and excess returns to the changes in credit spreads. I start from the simple case in which the state vector includes only excess returns, $\tilde{r}_{i,t}^e$, credit spreads, $d_{i,t}\tilde{s}_{i,t}$, and the probability of default in the Merton model times duration, $\tau\tilde{PD}_{i,t}$ ⁶. (I drop the subscripts from $\tau_{i,t}$ to save notation.) I use excess returns instead of credit loss, as credit loss in the right-hand side of the regression is mostly zero. I include probability of default based on the Merton (1974) model, because it is known to forecast default (e.g., Gropp, Loduca and Vesala (2006) and Harada, Ito and Takahashi (2010)), and Gilchrist and Zakrajšek (2012) use the Merton (1974) model to decompose their measure of credit spreads.

I run pooled OLS regressions using demeaned state variables to estimate the VARs. To account for the cross-sectional correlation in error terms, I cluster standard errors by time.⁷

Table 1 shows the summary statistics of the variables. The statistics are computed using the panel data of all bonds in the sample. Panel A shows the raw data before demeaning. The excess returns are distributed symmetrically, while the probability of default, credit spreads and credit loss are right-skewed. Panel B shows the demeaned data to for the VAR estimates, in which the cross-sectional mean is subtracted from each observation. Demeaning does not significantly reduce the volatility of the variables, while it somewhat reduces the skewness of credit loss.

Panel C presents the estimated VAR coefficients. Excess returns tend to be higher when

⁶ $PD_{i,t} = \Phi(-d_{2,i,t})$,
where

$$d_{2,i,t} = \frac{\log A_{i,t}/K + (rf - 0.5\sigma_A^2)}{\sigma_A}$$

and $A_{i,t}$ is the firm's asset value, K is the book value of short-term debt plus half of the long-term debt, rf is the risk-free rate and σ_A is the asset volatility. I use rf following Bharath and Shumway (2008). The methodology to compute $PD_{i,t}$ is explained in the online appendix. Using $d_{2,i,t}$ in place of the probability of default, $PD_{i,t}$, does not change the results.

⁷In the online appendix, I compare the clustered standard errors with the standard errors from bootstrapping which confirms the reliability of the statistical inference.

Table 1: Summary Statistics of the Variables and Estimated VAR: Monthly from 1973 to 2011

Panel A: Descriptive Statistics, Basic Data							
Variable	Mean	Std.	5%-pct	25%-pct	Median	75%-pct	95%-pct
$r_{i,t}^e$	0.07	3.22	-4.10	-0.96	0.11	1.18	4.24
$s_{i,t}$	11.11	11.09	1.27	4.14	8.04	14.62	30.31
$l_{i,t}$	0.03	2.10	0	0	0	0	0
$\tau PD_{i,t}$	2.01	15.47	0.00	0.00	0.00	0.00	2.89
$r_{i,t}^{eq}$	0.82	9.15	-12.52	-3.31	1.08	5.38	13.55
$bm_{i,t}$	-26.50	75.76	-160.07	-67.13	-13.96	21.86	68.80
Panel B: Descriptive Statistics, Demeaned Data							
Variable	Mean	Std.	5%-pct	25%-pct	Median	75%-pct	95%-pct
$\tilde{r}_{i,t}^e$	0	2.69	-2.84	-0.78	0.01	0.81	2.93
$\tilde{s}_{i,t}$	0	10.16	-10.92	-5.65	-2.04	3.51	16.71
$\tilde{l}_{i,t}$	0	2.09	-0.16	0	0	0	0
$\tau \tilde{P}D_{i,t}$	0	14.90	-6.74	-1.47	-0.64	-0.30	0.54
$\tilde{r}_{i,t}^{eq}$	0	7.94	-11.08	-3.57	0.04	3.75	11.19
$\tilde{b}m_{i,t}$	0	65.49	-108.83	-28.54	4.92	32.15	94.00
Panel C: VAR estimates, $A \times 100$							
	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t}d_{i,t}^{Baa}$	$\tilde{s}_{i,t}d_{i,t}^{Ba}$	$\tilde{s}_{i,t}d_{i,t}^{B-}$	$\tau \tilde{P}D_{i,t}$	R^2
$\tilde{r}_{i,t+1}^e$	-3.61 (2.02)	2.81 (0.67)	-0.16 (0.30)	0.24 (0.61)	-2.23 (0.72)	-0.31 (0.27)	0.01
$\tilde{s}_{i,t+1}$	10.51 (2.76)	97.76 (0.68)	0.29 (0.32)	-0.21 (0.62)	-3.03 (1.09)	0.28 (0.25)	0.91
$\tilde{s}_{i,t+1}d_{i,t+1}^{Baa}$	3.93 (0.47)	0.91 (0.12)	95.07 (0.81)	0.16 (0.26)	-0.68 (0.15)	-0.07 (0.10)	0.90
$\tilde{s}_{i,t+1}d_{i,t+1}^{Ba}$	1.97 (0.50)	0.18 (0.07)	1.31 (0.22)	92.49 (0.94)	0.11 (0.09)	-0.10 (0.06)	0.84
$\tilde{s}_{i,t+1}d_{i,t+1}^{B-}$	0.10 (2.23)	-0.16 (0.08)	0.23 (0.09)	3.68 (0.59)	93.83 (1.16)	0.61 (0.15)	0.86
$\tau \tilde{P}D_{i,t+1}$	-4.14 (1.41)	1.74 (0.39)	-0.36 (0.45)	1.62 (0.63)	3.29 (0.89)	96.69 (1.06)	0.93

Panel A reports the statistics for the raw data, while in Panel B the variables are market-adjusted by subtracting the cross-sectional average each month. Means, standard deviations and percentiles (5, 25, 50, 75, and 95 percent) are estimated using the monthly panel data from January 1973 to December 2011. All the variables are shown in percentage. $r_{i,t}^e$ is the log return on the corporate bonds in excess of the matching Treasury bond, $l_{i,t}$ is the credit loss, $s_{i,t}$ is the credit spread of the corporate bonds, $\tau PD_{i,t}$ is the probability of default implied by the Merton model times the bond's duration, $r_{i,t}^{eq}$ is issuer's equity return in excess of T-bill rate and $bm_{i,t}$ is log issuer's equity book-to-market ratio. Panel C shows the estimated VAR coefficients, multiplied by 100. $d_{i,t}^\theta$ is a dummy variable for the rating θ . The number of observations is 791,864 bond months, and there are 260 default observations.

past excess returns are low, credit spreads are high, or the issuer is less likely to default. The predictive power of credit spreads is strong for most of the credit ratings, with a coefficient of 2.81 for A+ (rated Aaa, Aa or A) bonds, $2.81-0.16=2.65$ for Baa bonds, and $2.81+0.24=3.05$ for Ba bonds. The credit spreads and probability of default are fairly autonomous and are forecastable mostly by their own past values.

Panel A of Table 2 shows the VAR-implied long-run forecasting coefficients in (8) and (9). Holding everything else constant, when credit spreads go up by one, the expected long-run credit loss goes up by 0.09 for A+ bonds, 0.18 for Baa bonds, 0.46 for Ba bonds and 0.89 for B- (rated B or below) bonds. Under the benchmark case of the expected credit loss hypothesis, the slope coefficient on credit spreads must be one. In the data, except for highly risky bonds, the estimated long-run credit loss forecasting coefficients are significantly below one. Panel A also shows that the probability of default in the Merton (1974) model helps forecast default in the long-run, with a coefficient estimate of 0.12.

Since credit spreads must forecast either credit loss or excess returns in the long-run, lower long-run credit loss forecasting coefficients imply higher return forecasting coefficients. A unit increase in credit spreads corresponds to an increase of 0.90 in risk premium for A+ bonds, 0.81 for Baa bonds, 0.54 for Ba bonds and 0.12 for B- bonds, showing significant dependence of the coefficients on ratings.

To examine the effect of the nonlinearity, I plot the long-run credit loss and excess return forecasting coefficients on credit spreads, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$ in Figure 1. Figure 1 visualizes how the slope differs across credit ratings and thereby shows the degree of nonlinearity in the long-run VAR. Within the range of IG ratings, the expected credit loss forecasting coefficients are close to zero, and thus the line is rather flat. In contrast, the excess return forecasting coefficients are close to one, leading to the steep line. This implies that the variation in credit spreads within the IG ratings corresponds mostly to the variation in expected excess returns. However, as the credit spread increases, the line for expected

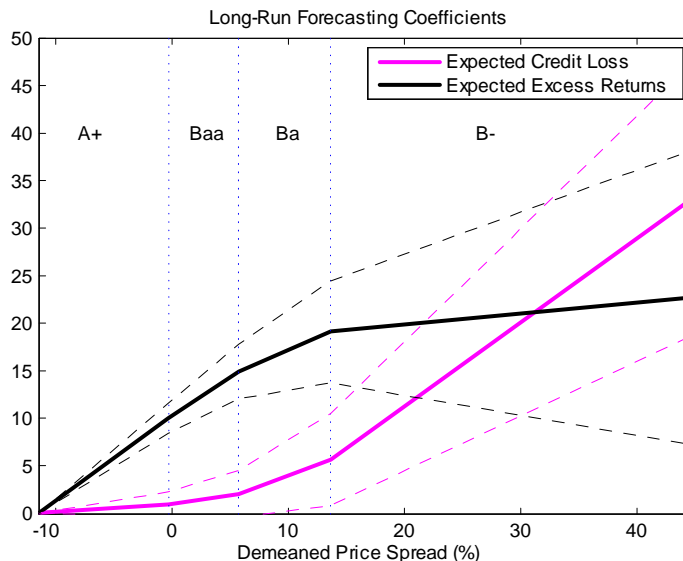
Table 2: Implied Long-Run Regression Coefficients and Volatility Ratios

Panel A: Long-run regression coefficients, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$							
	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t} d_{i,t}^{Baa}$	$\tilde{s}_{i,t} d_{i,t}^{Ba}$	$\tilde{s}_{i,t} d_{i,t}^{B-}$	$\tau \tilde{P} D_{i,t}$	$\sigma(E_t[\cdot])$
$\sum_{j=1}^{\bar{T}} \rho^{j-1} \tilde{l}_{i,t+j}$	-0.05 (0.02)	0.09 (0.06)	0.09 (0.04)	0.37 (0.09)	0.80 (0.09)	0.12 (0.06)	6.71
$\sum_{j=1}^{\bar{T}} \rho^{j-1} \tilde{r}_{i,t+j}^e$	0.05 (0.02)	0.90 (0.07)	-0.09 (0.04)	-0.36 (0.09)	-0.78 (0.09)	-0.12 (0.06)	5.25
Panel B: Variation of VAR-implied conditional expectations							
Estimates	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$		
	0.67 (0.12)	0.52 (0.06)	0.76 (0.03)	0.64 (0.14)	0.18 (0.20)		
Panel C: Regression of credit loss on information							
	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t} d_{i,t}^{Baa}$	$\tilde{s}_{i,t} d_{i,t}^{Ba}$	$\tilde{s}_{i,t} d_{i,t}^{B-}$	$\tau \tilde{P} D_{i,t}$	R^2
$\tilde{l}_{i,t+1}$	-6.98 (2.09)	0.11 (0.09)	-0.18 (0.07)	-0.06 (0.23)	5.59 (0.88)	0.02 (0.14)	0.05
$\tilde{l}'_{i,t+1}$	-6.82 (2.13)	0.19 (0.09)	-0.13 (0.08)	-0.04 (0.24)	5.24 (0.89)	0.03 (0.14)	0.04
Panel D: Directly forecasting credit loss							
Estimates	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$		
	0.69 (0.13)	0.54 (0.05)	0.75 (0.03)	0.60 (0.15)	0.10 (0.20)		
Panel E: Long VAR							
Estimates	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$		
	0.52 (0.11)	0.59 (0.06)	0.73 (0.05)	0.63 (0.16)	0.24 (0.16)		

The sample period is monthly from 1973 to 2011. Panel A shows the VAR-implied long-run coefficient for a bond with the average maturity, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$ where $G(\bar{T}) = A(I - \rho A)^{-1} (I - (\rho A)^{\bar{T}-t})$. $\sigma(E_t[\cdot])$ shows the sample standard deviation of fitted values of the left-hand side variables. $d_{i,t}^\theta$ is a dummy variable for the rating θ . Panel B shows the summary statistics of the long-run expected credit loss, $\tilde{s}_{i,t}^l = e_L G(T_i) X_{i,t}$, and the long-run expected returns, $\tilde{s}_{i,t}^r = e_1 G(T_i) X_{i,t}$. $\varrho(\cdot, \cdot)$ shows the sample correlation coefficient. Panel C shows the credit loss forecasting regression, $\tilde{l}_{i,t} = b_l X_{i,t} + \varepsilon_{i,t}^l$. $\tilde{l}'_{i,t}$ is the credit loss implied from the identity (1), so that $\tilde{l}'_{i,t} \equiv -\rho s_{i,t} + s_{i,t-1} - r_{i,t}^e$. Panel D shows the summary statistics of the long-run expected credit loss and excess returns, based on the VAR where I replace $\tilde{r}_{i,t+1}^e$ with $\tilde{r}_{i,t+1}^{e'}$ $\equiv -\rho \tilde{s}_{i,t+1} + \tilde{s}_{i,t} - \tilde{l}_{i,t+1}$. Panel E shows the estimates based on the VAR, where the state variables include lagged credit spreads times rating dummies, 3 lags of bond excess returns, probability of default, the issuers' stock returns, log book-to-market ratio, log market size of equity and log share price (winsorized at 15 dollars). Standard errors, reported in parentheses under each coefficient, are clustered by time.

credit loss starts to steepen, while the line for expected excess returns begins to flatten out.

Figure 1: Long-Run Forecasting Coefficients By Credit Ratings



The x-axis is the demeaned credit spreads, with the left end set by the 1st percentile of credit spreads and the right end set by the 99th percentile. The y-axis is the long-run expected credit loss, $E_t \left[\sum \rho^{j-1} \tilde{l}_{i,t+j} \right]$, and excess returns, $E_t \left[\sum \rho^{j-1} \tilde{r}_{i,t}^e \right]$, and the slope of the line is the long-run forecasting coefficients in Panel A of Table 2. Dashed lines denote \pm standard-error bounds. The borders between credit ratings are set where the histogram of the credit spreads for the credit rating overlaps with the histogram for the neighboring credit ratings.

Despite the low R-squared in the return forecasting regression in Panel C of Table 1, the return predictability is economically significant: The standard deviation of expected returns is 0.24 percent per month (not reported in the table) and 5.25 percent in the long-run (Panel A, Table 2). The variation in expected returns is large compared with the variation found in the previous literature. For example, Gebhardt, Hvidkjaer and Swaminathan (2005) find that the difference in average excess returns between different credit ratings is 0.07 percent per month and the difference between different durations is 0.04 percent.

Panel B of Table 2 shows the ratio of the volatility of expected credit loss and excess returns to the credit spreads, quantifying the magnitude of the contribution of these two components. The volatility ratio for the credit loss is 0.67, while the ratio for the risk premium is 0.52. Thus, the magnitude of variation of these two components of credit

spreads are comparable to each other. The correlation between these two components and credit spreads is also similar to each other at 0.76 and 0.64. The volatility ratio for expected excess returns is highly significantly different from zero, and thus the expected credit loss hypothesis is rejected in the data. The correlation between the expected credit loss and excess returns is positive but insignificant.

In this VAR, I forecast credit loss indirectly by forecasting returns and credit spreads. Whether forecasting credit loss directly or indirectly does not matter if the log-linear approximation in (1) holds well. In Panel C, I compare the credit loss forecasting regressions for $\tilde{l}_{i,t}$ and $\tilde{l}'_t \equiv -\rho\tilde{s}_{i,t+1} + \tilde{s}_{i,t} - \tilde{r}_{i,t+1}^e$ using the same state vector as Panel A. The regression coefficients for $\tilde{l}_{i,t}$ and \tilde{l}'_t are similar to each other, and the gaps are within one standard error. In Panel D, I report the variance decomposition results based on the VAR replacing $\tilde{r}_{i,t+1}^e$ with $\tilde{r}_{i,t+1}^{e'} \equiv -\rho\tilde{s}_{i,t+1} + \tilde{s}_{i,t} - \tilde{l}_{i,t+1}$, and thus forecasting credit loss directly. The volatility ratio for expected credit loss becomes 0.69, little changed from the estimate in Panel B (0.67). Thus, the approximation error is not driving the results, and it does not matter whether I forecast excess returns or credit loss.

Panel E of Table 2 shows the estimates based on a “long” VAR which adds 2 extra lags of excess returns and probability of default, and 3 lags of the issuers’ stock returns, log book-to-market ratio, log market size of equity and log share price (winsorized at 15 dollars) to the main VAR specification in (6). To select these state variables, I first run a VAR using all the state variables tested in Duffie, Saita and Wang (2007) and Campbell, Hilscher and Szilagyi (2008) in forecasting defaults, and choose the variables that remain significant in forecasting long-run credit loss in my sample. The resulting volatility ratio is 0.52 for the expected credit loss and 0.59 for the expected excess returns. The correlations between the two components and credit spreads are 0.74 and 0.65. Therefore, the overall results that the contributions of the two components to the variation in credit spreads are comparable to each other do not depend on a particular VAR specification.

In the online appendix, I show that the variance decomposition results are robust to the small sample biases, and not affected by the state tax effects pointed out by Elton, Gruber, Agrawal and Mann (2001). Though the state tax can affect the level of the state variables, it does not change their movements. Finally, I show that the main results in Table 2 are robust, even if I interact $r_{i,t}^e$ and $\tau PD_{i,t}$ with the rating dummies, interact all variables with duration dummies to account for the maturity effect non-parametrically, or include industry fixed effects in estimating the VAR to account for the difference in credit spreads across industries.

4 Joint Decomposition with Stocks

In this section, I show the joint variance decomposition of bonds and the book-to-market ratio of stocks, and examine the interaction between bonds and stocks. I work on the book-to-market ratio rather than the dividend price ratio, as I focus on the issue-level variation in stock prices and many firms don't pay dividends. Vuolteenaho (2002) shows that the log book-to-market ratio of a stock can be expressed using a present value identity,

$$bm_{i,t} \approx E \left[\sum_{j=1}^{\infty} \rho_{eq}^{j-1} r_{i,t+j}^{eq} \middle| \mathcal{F}_t \right] + E \left[\sum_{j=1}^{\infty} \rho_{eq}^{j-1} (y_{i,t+j} - r f_{t+j}) \middle| \mathcal{F}_t \right], \quad (11)$$

where $bm_{i,t}$ is the log book-to-market ratio, $r_{i,t+j}^{eq}$ is a return on stock in excess of the log T-bill rate, $y_{i,t+j}$ is a log book return on equity defined by $y_{i,t+j} = \log(1 + Y_{t+j}/B_{t+j-1})$ where Y_{t+j} is earnings and B_{t+j-1} is book equity, and $r f_{t+j}$ is log T-bill rates. The discount coefficient ρ_{eq} is set to be 0.967, following Vuolteenaho (2002).

Equation (11) shows that a stock's book-to-market ratio can be decomposed into the risk premium and profitability components. By jointly decomposing bonds and stocks, we can study the interaction in risk premiums and cash flows between bonds and stocks. If the Merton (1974) model holds, then the risk premiums and cash flows for bonds and stocks are

perfectly correlated. If there are more risk factors other than firm value, or if the bond and stock markets are segmented, then such relationship may break down.

I estimate the conditional expectations by jointly estimating the VAR for bonds and stocks. To this end, I augment the state vector with stock variables,

$$X_{i,t} = \begin{pmatrix} \tilde{r}_{i,t}^e & d_{i,t}\tilde{s}_{i,t} & \tau\tilde{P}D_{i,t} & \tilde{r}_{i,t}^{eq} & \tilde{b}m_{i,t} \end{pmatrix}, \quad (12)$$

which follows the same dynamics $X_{i,t} = AX_{i,t-1} + W_{i,t}$. Then the long-run risk premium on the stock can be found by

$$\begin{aligned} \tilde{b}m_{i,t}^y &\equiv E \left[\sum_{j=1}^{\infty} \rho_{eq}^{j-1} (\tilde{y}_{i,t+j} - rf_{i,t+j}) \middle| X_{i,t} \right] = e_y G_{eq}(\infty) X_{i,t}, \\ \tilde{b}m_{i,t}^r &\equiv E \left[\sum_{j=1}^{\infty} \rho_{eq}^{j-1} \tilde{r}_{i,t+j}^{eq} \middle| X_{i,t} \right] = e_7 G_{eq}(\infty) X_{i,t}, \end{aligned}$$

where $G_{eq}(\infty) = A(I - \rho_{eq}A)^{-1}$, $e_y = e_7 + \rho_{eq}e_8 - e_8A^{-1}$, e_7 is a unit vector whose seventh entry (corresponding to $\tilde{r}_{i,t}^{eq}$) is one and other entries are zero, and e_8 is a unit vector whose eighth entry (corresponding to $\tilde{b}m_{i,t}$) is one and other entries are zero.

As I use a subsample of firms who issue corporate bonds, the stocks in my analysis are quite different from the entire universe of stocks. Most notably, 84 percent of the observations (in bond-months) correspond to “Big” stocks that are larger than the 50th NYSE percentile, 16 percent is “Small” stocks that are between the 20th and 50th NYSE percentiles, while the fraction for “Micro” stocks is negligible. In contrast, Fama and French (2008) report that “Micro” stocks account for more than half of their sample of stocks. Large firms issue more bonds than small firms do, and thus my sample tends to be dominated by large firms who have multiple corporate bond issues.

I report the estimated VAR coefficients in Appendix C. Table 3 reports the VAR-implied long-run expectations. Panels A and B show that including stock variables does not ma-

Table 3: Joint Decomposition of Bond and Stock Prices

Panel A: Long-run regression coefficients, $e_L G(T)$, $e_1 G(T)$, $e_y G_{eq}(\infty)$ and $e_7 G_{eq}(\infty)$									
	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t} d_{i,t}^{Baa}$	$\tilde{s}_{i,t} d_{i,t}^{Ba}$	$\tilde{s}_{i,t} d_{i,t}^{B-}$	$\tau \tilde{P} D_{i,t}$	$\tilde{r}_{i,t}^{eq}$	$\tilde{b} m_{i,t}$	$\sigma(E_t[\cdot])$
$\sum_{j=1}^{\bar{T}} \rho^{j-1} \tilde{l}_{i,t+j}$	-0.04 (0.01)	0.04 (0.04)	0.10 (0.04)	0.35 (0.10)	0.78 (0.13)	0.06 (0.04)	-0.03 (0.01)	0.02 (0.01)	5.96
$\sum_{j=1}^{\bar{T}} \rho^{j-1} \tilde{r}_{i,t+j}^e$	0.04 (0.01)	0.94 (0.05)	-0.09 (0.04)	-0.34 (0.10)	-0.77 (0.13)	-0.06 (0.04)	0.03 (0.01)	-0.02 (0.01)	5.48
$\sum_{j=1}^{\infty} \rho^{j-1} (\tilde{y}_{i,t+j} - r f_{t+j})$	0.11 (0.04)	0.01 (0.13)	0.03 (0.09)	-0.33 (0.17)	-0.76 (0.21)	-0.19 (0.15)	0.02 (0.02)	-0.99 (0.03)	66.73
$\sum_{j=1}^{\infty} \rho^{j-1} \tilde{r}_{i,t+j}^{eq}$	0.11 (0.04)	0.01 (0.13)	0.03 (0.09)	-0.33 (0.17)	-0.76 (0.21)	-0.19 (0.15)	0.02 (0.02)	0.01 (0.03)	6.52

Panel B: Variation of VAR-implied conditional expectations for corporate bonds

	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$
Estimates	0.60 (0.13)	0.55 (0.06)	0.75 (0.02)	0.69 (0.13)	0.25 (0.20)

Panel C: Variance decomposition of stocks and their correlation

	$\frac{\sigma(\tilde{b} m^y)}{\sigma(\tilde{b} m)}$	$\frac{\sigma(\tilde{b} m^r)}{\sigma(\tilde{b} m)}$	$\varrho(\tilde{s}^l, \tilde{b} m^y)$	$\varrho(\tilde{s}^r, \tilde{b} m^r)$	$\varrho(\tilde{s}^l, \tilde{b} m^r)$	$\varrho(\tilde{s}^r, \tilde{b} m^y)$
Estimates	1.02 (0.04)	0.10 (0.04)	-0.43 (0.06)	-0.16 (0.18)	-0.94 (0.07)	0.03 (0.13)

The sample period is monthly from 1973 to 2011. Panel A shows the VAR-implied long-run coefficient for a bond with the average maturity, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$ where $G(\bar{T}) = A(I - \rho A)^{-1} (I - (\rho A)^{\bar{T}-t})$, and for stocks, $e_y G_{eq}(\infty)$ and $e_7 G_{eq}(\infty)$, where $G_{eq}(\infty) = A(I - \rho_{eq} A)^{-1}$. $\tilde{r}_{i,t}^{eq}$ is log equity return in excess of T-bill rate, $\tilde{b} m_{i,t}$ is log book-to-market ratio, $\tilde{y}_{i,t}$ is book return on equity, $r f_t$ is T-bill rate, $d_{i,t}^\theta$ is a dummy variable for the rating θ . $\sigma(E_t[\cdot])$ shows the sample standard deviation of fitted values of the left-hand side variables. Panel B shows the summary statistics of the long-run expected credit loss for bonds, $\tilde{s}_{i,t}^l = e_L G(T_i) X_{i,t}$, and the long-run expected returns, $\tilde{s}_{i,t}^r = e_1 G(T_i) X_{i,t}$. $\varrho(\cdot, \cdot)$ shows the sample correlation coefficient. Panel C shows the summary statistics for the long-run expected profitability for stocks, $\tilde{b} m_{i,t}^y = e_y G_{eq}(\infty) X_{i,t}$, and the long-run stock risk premiums, $\tilde{b} m_{i,t}^r = e_7 G_{eq}(\infty) X_{i,t}$, where $G_{eq}(\infty) = A(I - \rho_{eq} A)^{-1}$. Standard errors, reported in parentheses under each coefficient, are clustered by time.

terially change the decomposition results for bonds. The expected returns and credit loss components each accounts for slightly more than half of the total variation in credit spreads.

Panel C shows the decomposition results for stocks and their relationships with the bond decomposition. The volatility ratio for the profitability, $\frac{\sigma(\tilde{b}m^y)}{\sigma(\tilde{b}m)}$, is 1.02, while the volatility ratio for the risk premium, $\frac{\sigma(\tilde{b}m^r)}{\sigma(\tilde{b}m)}$, is only 0.10. These findings are consistent with Vuolteenaho (2002), who finds that for large stocks, the cash flow component is the major source of variation of the book-to-market ratio. Specifically, Vuolteenaho (2002) reports that for big stocks, the volatility of discount rate shocks is $\sqrt{0.0040} = 0.06$, while the volatility of cash flow shocks is $\sqrt{0.0319} = 0.18$ (see his Table 4). Since my sample is tilted toward large firms, much of the variation in stock prices corresponds to the expected profitability variation.

Panel C also shows that the correlation between bonds' expected credit loss and stocks' profitability, $\rho(\tilde{s}^l, \tilde{b}m^y)$, is negative and statistically significant. This estimate seems reasonable, as more profitable firms are less likely to default. In contrast, the correlation in risk premium between bonds and stocks, $\rho(\tilde{s}^r, \tilde{b}m^r)$, is insignificant. If the Merton (1974) model holds and leverage is (cross sectionally) constant, then the correlation in risk premium should be one. The weak risk premium correlation suggests that the variation in leverage or risk-factors other than firms' asset value breaks the correlation between bond and stock risk premiums.

Interestingly, the correlation between the expected credit loss for bonds and the risk premiums for stocks, $\rho(\tilde{s}^l, \tilde{b}m^r)$, is statistically significantly negative. The firms closer to default earns lower expected returns on their stocks. This negative correlation is consistent with Campbell, Hilscher and Szilagyi (2008), who find the distress anomaly using the entire universe of stocks. However, the negative correlation poses a challenge for the rational asset pricing models which typically predict that riskier stocks earn higher expected returns.

To better understand the distress anomaly, we turn to Panel A of Table 3, which shows

the long-run forecasting coefficients for bonds and stocks. Throughout the credit ratings, holding everything else constant, higher spreads forecast higher expected credit loss for bonds, and the effect is more pronounced for high yield bonds. In contrast, higher credit spreads insignificantly predict long-run stock returns for IG bonds, while higher spreads predict lower stock returns for high yield bond issuers. Once we control for credit spreads, log book-to-market ratio does not help predict returns on bonds or stocks. As a result, I find the distress anomaly in which stocks with higher default risk earn lower expected returns.

5 Aggregate Credit Spread Dynamics and the Effects on Investment

5.1 Decomposition of the Market Portfolio of Corporate Bonds

Campbell and Shiller (1988a and 1988b) and Cochrane (2008 and 2011) emphasize the importance of time-varying risk premiums in understanding the price of the stock market portfolio. In contrast, Vuolteenaho (2002) finds that cash flow shocks are more important for individual stocks. Thus far, I find that the expected default component is about as important as the expected excess return component for individual corporate bonds. However, given the evidence in the stock market, these results may be different for the aggregate corporate bond market portfolio. To examine the difference for the aggregate market, I take the equal-weighted average of individual variables in each month to obtain the aggregate variables, and denote them with subscripts EW . For example, the equal-weighted market portfolio returns are computed by

$$r_{EW,t}^e \equiv \frac{1}{N_t} \sum_{i=1}^{N_t} r_{i,t}^e$$

where N_t is the number of bonds in month t . These equal-weighted average returns and credit spreads are an approximation to the logarithm of the market returns and spreads, as the average of the logarithm is not, in general, equal to the logarithm of the averages.

Using these aggregate variables, I run a restricted VAR with a state vector:

$$X_{i,t} = \begin{pmatrix} r_{i,t}^e & d_{i,t}s_{i,t} & \tau PD_{i,t} & r_{EW,t}^e & s_{EW,t} & \tau PD_{EW,t} \end{pmatrix}, \quad (13)$$

which follows the dynamics

$$X_{i,t+1} = AX_{i,t} + W_{i,t+1}.$$

I restrict the three-by-six entries at the lower left corner of the matrix A to be zero, so that the current individual variables do not forecast the future aggregate variables. By including the aggregate variables, I can exploit the cross-sectional variation of individual bonds without demeaning. Thus, based on this VAR with aggregate variables, the cross-sectional average of the estimated expected credit loss and excess returns will not be zero, making it possible to examine the variation in the average expected credit loss and excess returns over time.

I show the estimated VAR coefficients in Appendix C. Panels A and B of Table 4 show that the variance decomposition for individual bonds do not change much after including aggregate variables in the VAR. For the cross-section of individual bonds, the volatility ratio for expected credit loss is 0.69, which is similar to the ratio for expected excess return (0.63).

Panel C shows the variance decomposition for the equal-weighted market portfolio, implied by the VAR. The difference in volatility ratios between the individual bond level and the aggregate level is large: At the aggregate portfolio level, the volatility ratio for the expected credit loss is only 0.27, while the ratio for the expected excess returns is 0.96, much higher than the expected credit loss. Indeed, nearly 100% of the time-series variation in

Table 4: Decomposition of the Equal-Weighted Market Portfolio

Panel A: Long-run regression coefficients, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$										
	$r_{i,t}^e$	$s_{i,t}$	$s_{i,t} d_{i,t}^{Baa}$	$s_{i,t} d_{i,t}^{Ba}$	$s_{i,t} d_{i,t}^{B-}$	$\tau PD_{i,t}$	$r_{EW,t}^e$	$s_{EW,t}$	$\tau PD_{EW,t}$	$\sigma(E_t[\cdot])$
$\sum_{j=1}^{\bar{T}} \rho^{j-1} l_{i,t+j}$	-0.06 (0.02)	0.07 (0.06)	0.15 (0.05)	0.41 (0.13)	0.76 (0.14)	0.10 (0.05)	0.07 (0.02)	-0.20 (0.12)	-0.27 (0.11)	7.91
$\sum_{j=1}^{\bar{T}} \rho^{j-1} r_{i,t+j}^e$	0.06 (0.02)	0.93 (0.06)	-0.16 (0.06)	-0.41 (0.13)	-0.76 (0.14)	-0.10 (0.05)	-0.07 (0.02)	0.21 (0.13)	0.27 (0.11)	7.19

Panel B: Variation of VAR-implied conditional expectations (individual bonds)

	$\frac{\sigma(s^l)}{\sigma(s)}$	$\frac{\sigma(s^r)}{\sigma(s)}$	$\varrho(s^l, s)$	$\varrho(s^r, s)$	$\varrho(s^l, s^r)$
Estimates	0.69 (0.14)	0.63 (0.04)	0.74 (0.03)	0.66 (0.17)	-0.01 (0.23)

Panel C: Variance decomposition of the aggregate portfolio

	$\frac{\sigma(s_{EW}^l)}{\sigma(s_{EW})}$	$\frac{\sigma(s_{EW}^r)}{\sigma(s_{EW})}$	$\varrho(s_{EW}^l, s_{EW})$	$\varrho(s_{EW}^r, s_{EW})$	$\varrho(s_{EW}^l, s_{EW}^r)$	Div. factor	
Estimates	0.27 (0.06)	0.96 (0.13)	0.56 (0.42)	0.97 (0.02)	0.36 (0.43)	s_{EW}^l	s_{EW}^r
						0.08	1.17

The sample period is monthly from 1973 to 2011. Panel A shows the VAR-implied long-run coefficient for a bond with the average maturity, $e_L G(\bar{T})$ and $e_1 G(\bar{T})$ where $G(\bar{T}) = A(I - \rho A)^{-1} (I - (\rho A)^{\bar{T}-t})$. $d_{i,t}^\theta$ is a dummy variable for the rating θ and $\sigma(E_t[\cdot])$ shows the sample standard deviation of fitted values of the left-hand side variables. The variables with subscript EW are the equal-weighted average across bonds, computed every month. Panel B shows the summary statistics of the long-run expected credit loss for individual bonds, $s_{i,t}^l = e_L G(T_i) X_{i,t}$, and the long-run expected returns, $s_{i,t}^r = e_1 G(T_i) X_{i,t}$. $\varrho(\cdot, \cdot)$ shows the sample correlation coefficient. Panel C shows the summary statistics for the aggregate expected credit loss, $s_{EW}^l = \frac{1}{N_t} \sum s_{i,t}^l$ and the aggregate risk premium, $s_{EW}^r = \frac{1}{N_t} \sum s_{i,t}^r$. Standard errors, reported in parentheses under each coefficient, are clustered by time. Div. factor is $\sigma^2(s_{EW,t}^u) / \bar{\sigma}^2(s_{i,t}^u)$, where $\bar{\sigma}^2(s_{i,t}^u)$ is the time-series variance for bond i , $\sigma^2(s_{i,t}^u)$, averaged across bonds.

the aggregate credit spreads corresponds to the variation in risk premiums. The correlation between the credit spread and risk premiums is close to one as well.

The large discrepancy in decomposition results between the cross-section of individual bonds and the aggregate portfolio is due to the diversification effects. Following Vuolteenaho (2002), I compute the diversification factor:

$$\text{Diversification Factor} = \frac{\sigma^2(s_{EW,t}^u)}{\bar{\sigma}^2(s_{i,t}^u)} \quad u \in \{l, r\},$$

where $\bar{\sigma}^2(s_{i,t}^u) \equiv \frac{1}{N} \sum_{i=1}^N \sigma_i^2(s_{i,t}^u)$. The diversification factor compares the variance of the market variable with the average of the variance of the individual variable. If the variation of the individual variable is idiosyncratic, then the diversification factor becomes close to zero. In contrast, if much of the variation of the individual variable comes from a systematic shock, then the diversification factor becomes larger. Table 4 shows that the diversification factor is 0.08 for the expected credit loss while it is 1.17 for the expected excess returns. The diversification factors show that much of the variation in individual bonds' expected credit loss is due to idiosyncratic shocks, while much of the variation in expected excess returns is from systematic shocks. Thus, my findings are consistent with the previous findings in stocks, in which the risk premium variation dominates the aggregate dynamics, while the cash flow variation is significant for individual securities.

5.2 Effects of Credit Spreads on Investment

In this section, I examine how the two components of credit spreads affect firms' investment in the future. Philippon (2009) shows that credit spreads can be a proxy for Tobin's q , and Gilchrist and Zakrajšek (2012) find that many macro economic variables are forecastable mainly by the "excess bond premium", or the residuals of credit spreads unexplained by the default risk based on the Merton (1974) model.

Tobin's q theory does not discriminate between the risk premium variation and cash flow variation as a determinant of investment. A firm should change its investment in response to a changing market value of its assets, regardless of whether the change comes from the risk premium or cash flow shocks. However, there are several reasons why different components of credit spreads may affect investments differently.

First, Stein (1996) argues that, if investors are irrational and a firm manager is rational, then the manager whose firm is not financially constrained should ignore the risk premium variation in making an investment decision to maximize the long-run firm value⁸. If firm managers follow this advice in reality, then the variation in the expected excess returns should not predict investments, while the expected credit loss component should. Second, the expected default component can affect firms' investment decisions due to managerial frictions and market segmentations (e.g. debt overhang of Myers (1977)). If the firm is close to default, the conflict between bond holders and equity holders intensifies, which can reduce the firm's investment. Third, Tobin's (average) q is a convex function of credit spreads (see Philippon's (2009) Figure 1) due to changing put option delta. A unit change in credit spreads for low spread bonds (which are mostly driven by risk premiums) corresponds to a larger change in Tobin's q than a change in spreads for high yield bonds (which are mostly associated with expected credit loss). Thus, the risk premium component may affect investments more than the expected default. As the above explanations work in opposite directions, which part of the credit spreads better forecasts investment is unclear based on the existing theories. Thus, I let the data tell which effects seem to dominate the other.

To this end, I use the decomposition results based on the VAR including macro variables in Table 4, in order to contrast the results at the individual firm level with the results at the aggregate level. I take the average of all the bonds issued by a firm to estimate the firm-level expected excess returns and credit loss. I forecast the investment rate, measured by the

⁸Consistent with this view of the financial market, Greenwood and Hanson (2013) find some evidence for time-varying mispricing in the corporate bond market.

ratio of the capital expenditures this fiscal year to capital (measured by property, plant and equipment) at the end of the previous year, using the variables as of the end of the previous year. Since firms' fiscal years end in different months, I use monthly bond data to find the exact fiscal year end and discard the bond information in other months of the fiscal year. I also use the book-to-market ratio, profitability, sales-to-capital ratio, idiosyncratic volatility and lagged investment rate of the firm as controls, following Gilchrist, Sim and Zakrajšek (2013). For this exercise, I exclude financial firms (SIC codes from 6000 to 6800), as the nature of investment and capital is different for the financial and nonfinancial industries.

First, I focus on the individual firm-level variation and run pooled OLS regressions. I demean all variables using the cross-sectional average every month, and exploit the cross-sectional variation. Panel A of Table 5 shows the results of the forecasting regressions. The first two columns show the forecasting regressions using each component of credit spreads separately, controlling only for the lagged value of the investment rate. Both the expected credit loss and excess return components negatively forecast the investment next period, with the estimated slope coefficients of -0.37 and -0.49, respectively.

I include all the other control variables in the next two columns. The forecasting power of the expected credit loss and risk premium decreases slightly, but it is still statistically significant. When the expected credit loss rises by one percent, the investment rate falls by 0.24 percent next year, while a one-percent rise in the expected excess returns leads to a 0.20 percent decrease in the investment rate. The results are similar when I include both the expected credit loss and excess returns in the same regression, as shown in the last column. Thus, at the firm level, both the expected default and risk premium components play a significant role in affecting individual firms' investment decisions.

To examine the effects at the aggregate level, I also take the equal-weighted average of all variables every year to obtain the market-level variable and run time-series regressions. Panel B of Table 5 shows the estimated coefficients of the forecasting regressions at the

Table 5: Investment Forecasting Regressions: Firm-Level Annual Data From 1973 to 2012

	Panel A: Individual Firms					Panel B: Equal-Weighted Market Portfolios				
	Left-hand side variable: $\log I/K_{k,t+12}$					Left-hand side variable: $\log I/K_{EW,t+12}$				
$E_t[\sum_j \rho^{j-1} l_{t+j}]$	-0.37 (0.08)		-0.24 (0.09)		-0.26 (0.09)	-3.36 (1.66)		-3.81 (1.24)		-0.27 (0.56)
$E_t[\sum_j \rho^{j-1} r_{t+j}^e]$		-0.49 (0.11)		-0.20 (0.09)	-0.29 (0.10)		-2.20 (0.22)		-2.74 (0.25)	-2.70 (0.23)
bm_t			-0.15 (0.01)	-0.15 (0.01)	-0.15 (0.01)			-0.19 (0.04)	0.02 (0.03)	0.02 (0.04)
$\log \Pi/K_t$			0.00 (0.15)	0.01 (0.16)	0.01 (0.15)			4.08 (1.47)	-1.36 (0.70)	-1.35 (0.72)
$\log Y/K_t$			0.06 (0.01)	0.06 (0.01)	0.06 (0.01)			0.44 (0.13)	0.35 (0.12)	0.35 (0.12)
σ_t^{IV}			0.03 (0.02)	0.02 (0.01)	0.03 (0.02)			-0.02 (0.04)	0.02 (0.03)	0.02 (0.03)
$\log I/K_t$	0.67 (0.02)	0.67 (0.02)	0.58 (0.02)	0.58 (0.02)	0.58 (0.02)	0.53 (0.10)	0.42 (0.07)	0.36 (0.08)	0.30 (0.07)	0.30 (0.07)
\bar{R}^2	0.48	0.48	0.51	0.51	0.51	0.32	0.75	0.54	0.82	0.82

Panel A shows the result of the forecasting regression of the log investment rate for firm k over the period between t and $t+12$, $\log I/K_{k,t+12}$, using the components of credit spreads in month t . The forecasting coefficients are estimated using pooled OLS regressions. The variables $E_t[\sum_j \rho^{j-1} l_{k,t+j}]$ and $E_t[\sum_j \rho^{j-1} r_{k,t+j}^e]$ are the long-run expected credit loss and excess returns estimated based on the VAR with the state vector in (13). The components of credit spreads for firm k are computed by taking the average of all bonds issued by firm k each month. The variable $\log \Pi/K_{k,t}$ is the log profitability (operating profit divided by property, plant and equipment), $\log Y/K_{k,t}$ is the log ratio of sales to capital, $\log \sigma_{k,t}^{IV}$ is the log idiosyncratic volatility computed following Ang, Hodrick, Xing and Zhang (2006), and $\log I/K_{k,t}$ is the lagged log investment rate for firm k in the fiscal year ending in month t . \bar{R}^2 is an adjusted R-squared. Standard errors, reported in parentheses, are clustered by time and adjusted for autocorrelation with Newey-West 12 lags. All variables are demeaned using the equal-weighted average every month, and winsorized at the 0.1th and 99.9th percentile. The number of observations is 7,300 firm years.

Panel B shows the result of the forecasting regression of the equal-weighted average of the log investment rate from t to $t+12$, $\log I/K_{EW,t+12}$. All explanatory variables are also the equal-weighted average of the individual firms. Standard errors, reported in parentheses, are adjusted for autocorrelation with Newey-West 3 lags. The number of observations is 39 years.

Figure 2: Equal-Weighted Average Investment Ratio, Expected Credit Loss and Expected Excess Returns

The figure plots the equal-weighted average investment ratio, $\log I/K_{EW,t}$, expected credit loss, $s_{EW,t}^l$, and expected excess returns, $s_{EW,t}^r$. The expected credit loss and excess returns are moved forward by one year, so I plot $s_{EW,t}^l$ and $s_{EW,t}^r$ in $t + 1$. All variables are demeaned.

market level. When the expected credit loss and excess returns are used separately, with or without other control variables, both of them negatively predict the investment rate next year. However, when the two components are put together in the same regression, the expected return dominates the expected credit loss in forecasting investment. Figure 2 shows the time series of the aggregate investment rate, expected credit loss and expected excess returns. The negative correlation between the investment rate and the expected excess return component of credit spreads (forwarded one year) is evident throughout the sample period. In contrast, the expected credit loss component moves little over time and is less correlated with the investment rate. Thus, my findings are consistent with the interpretation that much of the management's reaction to bonds' mispricing and frictions among the firm's stake holders, if they exist, are firm-specific, and affect only the individual firms' investment decision, not the aggregate investment.

6 Conclusion

I show that the credit spreads of corporate bonds can be decomposed into an expected excess return component and an expected credit loss component without relying on a particular model of default. Applying the Campbell-Shiller (1988a) style decomposition, I show that about half of the cross-sectional variation of the credit spreads corresponds to changes in the risk premium, and its volatility is almost as large as that of the expected credit loss.

By estimating the VARs including market-level variables, I contrast the firm- or bond-level results with the decomposition of the market portfolio. Though the expected credit

loss is as important as the expected excess returns at the individual bond level, the risk premium component is the dominating factor in the aggregate credit spread dynamics. Since much of the expected default loss variation at the security level is idiosyncratic, the credit loss components are mostly diversified away in the aggregate market, and their aggregate volatility is small.

In addition to forecasting corporate investment, understanding the information in credit spreads is important for a dynamic portfolio choice problem, since part of the variation in credit spreads signals the variation in expected returns. The decomposition is also important for credit risk management, as one might use credit spreads to measure default risk. My analysis shows that credit spreads forecast both excess returns and default in the future, and thus provide a useful signal for portfolio management.

One analysis left for the future is to explore the role of illiquidity in corporate bonds within the variance decomposition framework. If an investor expects the corporate bond will become illiquid when she has to sell in the future, then she might discount the valuation of the bond today, leading to a variation in credit spreads. The variance decomposition approach can be easily extended to account for illiquidity, though the empirical measurement of illiquidity poses a challenge for the extension.

References

- [1] Ang, Andrew, Robert J. Hodrick, Yuhang Xing and Xiaoyan Zhang, 2006, The Cross-Section of Volatility and Expected Returns, *Journal of Finance* 61, 259-299.
- [2] Beber, Alessandro, Brandt, Michael W. and Kavajecz, Kenneth A., 2009, Flight-to-Quality or Flight-to-Liquidity? Evidence from the Euro-Area Bond Market, *Review of Financial Studies* 22: 925-957.
- [3] Bhamra, Harjoat S., Lars-Alexander Kuehn and Ilya A. Strebulaev, 2010, The Levered Equity Risk Premium and Credit Spreads: A Unified Framework, *Review of Financial Studies* 23, 2, 645-703.
- [4] Bharath, Sreedhar T. and Tyler Shumway, 2008, Forecasting Default with the Merton Distance to Default Model, *Review of Financial Studies* 21, 3, 1339-1369.
- [5] Bongaerts, Dion, 2010, Overrated Credit Risk: Three Essays on Credit Risk in Turbulent Times, *Working Paper*.
- [6] Campbell, John Y., 1991, A Variance Decomposition for Stock Returns, *Economic Journal* 101, 405, 157-179.
- [7] Campbell, John Y., Jens Hilscher and Jan Szilagyi, 2008, In Search of Distress Risk, *Journal of Finance* 63, 6, 2899-2939.
- [8] Campbell, John Y., and Robert J. Shiller, 1988a, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.
- [9] Campbell, John Y., and Robert J. Shiller, 1988b, Stock Prices, Earnings, and Expected Dividends, *Journal of Finance* 43, 3, 661-676.
- [10] Chen, Hui, 2010, Macroeconomic Conditions and the Puzzles of Credit Spreads and Capital Structure, *Journal of Finance* 65, 6, 2171-2212.
- [11] Chen, Long, David A. Lesmond, Jason Wei, 2007, Corporate Yield Spreads and Bond Liquidity, *Journal of Finance* 62, 1, 119-149.
- [12] Chen, Long, 2009, On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics* 92, 128-151.
- [13] Chen, Long, Pierre Collin-Dufresne and Robert S. Goldstein, 2009, On the Relation Between the Credit Spread and the Equity Premium Puzzle, *Review of Financial Studies* 22 (9), 3367-3409.
- [14] Chen, Nai-Fu, Richard Roll and Stephan A. Ross, 1986, Economic Forces and the Stock Market, *Journal of Business* 59, 383-403.
- [15] Cochrane, John H., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533-1575.

- [16] Cochrane, John H., 2011, Discount rates, *Journal of Finance* 66, 1047-1108.
- [17] Collin-Dufresne, Pierre, Robert S. Goldstein, 2001, Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* 56, 1929-1957.
- [18] Collin-Dufresne, Pierre, Robert S. Goldstein and J. Spencer Martin, 2001, The Determinants of Credit Spread Changes, *Journal of Finance* 56, 6, 2177-2207.
- [19] Crabbe, Leland E., 1991, Callable corporate bonds: A vanishing breed, *FEDS working paper #155*, Board of Governors of the Federal Reserve System.
- [20] Driessen, Joost, 2005, Is Default Event Risk Priced in Corporate Bonds?, *Review of Financial Studies* 18, 1, 165-195.
- [21] Duffee, Gregory R., 1999, Estimating the Price of Default Risk, *Review of Financial Studies* 12, 197-226.
- [22] Duffie, Darrell , Leandro Saita and Ke Wang, 2007, Multi-Period Corporate Default Prediction with Stochastic Covariates, *Journal of Financial Economics* 83, 635-665.
- [23] Duffie, Darrell and Kenneth J. Singleton, 1999, Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies* 12, 687-720.
- [24] Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, Christopher Mann, 2001, Explaining the Rate Spread on Corporate Bonds, *Journal of Finance* 56, 1, 247-277.
- [25] Fama, Eugene F., 1984, Forward and Spot Exchange Rates, *Journal of Monetary Economics* 14, 319-38.
- [26] Fama, Eugene F. and Robert R. Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* 77, 4, 680-692.
- [27] Gebhardt, William R., Soeren Hvidkjaer and Bhaskaran Swaminathan, 2005, The Cross-Section of Expected Bond Returns: Beta or Characteristics?, *Journal of Financial Economics* 75, 85-114.
- [28] Gilchrist, Simon, Vladimir Yankov and Egon Zakrajšek, 2009, Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets, *Journal of Monetary Economics* 56, 471-493.
- [29] Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit Spreads and Business Cycle Fluctuations, *American Economic Review* 102, 4, 1692-1720.
- [30] Gilchrist, Simon, Jae W. Sim and Egon Zakrajšek, 2013, Uncertainty, Financial Frictions and Irreversible Investment, *Working Paper*.
- [31] Greenwood, Robin and Samuel G. Hanson, 2013, Issuer Quality and Corporate Bond Returns, *Review of Financial Studies* 26, 1483-1525.

- [32] Gropp, Reint, Jukka Vesala and Giuseppe Vulpes, 2006, Equity and bond market signals as leading indicators of bank fragility, *Journal of Money, Credit and Banking* 38, 399-428.
- [33] Harada, Kimie, Takatoshi Ito and Shuhei Takahashi, 2010, Is the distance to default a good measure in predicting bank failures? Case studies, *Working paper*.
- [34] Hansen, Lars. P. and Robert J. Hodrick, 1980, Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis, *Journal of Political Economy* 88, 829–53.
- [35] Huang, Jingzhi and Ming Huang, 2012, How Much of the Corporate-Treasury Yield Spread is Due to Credit Risk?, *Review of Asset Pricing Studies* 2, 153-202.
- [36] Leland, Hayne E., 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, *Journal of Finance* 49, 4, 1213-1252.
- [37] Longstaff, Francis, Sanjay Mithal and Eric Neis, 2005, Corporate yield spreads: Default risk or liquidity? New evidence from the credit-default swap market, *Journal of Finance* 60, 2213-2253.
- [38] Lustig, Hanno, and Adrien Verdelhan, 2007, The Cross-section of Foreign Currency Risk Premia and Consumption Growth Risk, *American Economic Review* 97, 89–117.
- [39] McAndrews, James, Asani Sarkar and Zhenyu Wang, 2008, The Effect of the Term Auction Facility on the London Inter-Bank Offered Rate, *Federal Reserve Bank of New York Staff Report* 335.
- [40] Merton, Robert C., 1974, On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance* 29, 449-470.
- [41] Moody's, 1999, Historical Default Rates of Corporate Bond Issuers, 1920-1998.
- [42] Moody's, 2011, Corporate Default and Recovery Rates, 1920-2010.
- [43] Myers, Stewart C., 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics* 5, 147–175.
- [44] Philippon, Thomas, 2009, The bond market's Q, *Quarterly Journal of Economics* 124, 3, 1011-1056.
- [45] Schwartz, Krista, 2013, Mind the Gap: Disentangling Credit and Liquidity in Risk Spreads, *Working Paper*.
- [46] Stein, Jeremy C., 1996, Rational capital budgeting in an irrational world, *Journal of Business* 69, 429-455.
- [47] Taylor, John B. and John C. Williams, 2009, A Black Swan in the Money Market, *American Economic Journal, Macroeconomics*, 1: 58-83.

- [48] Vuolteenaho, Tuomo, 2002, What Drives Firm-Level Stock Returns?, *Journal of Finance* 57, 1, 233-264.
- [49] Warga, Arthur and Ivo Welch, 1993, Bondholder Losses in Leveraged Buyouts, *Review of Financial Studies* 6, 959-982.

A Derivation of the Credit Spread Decomposition

In this appendix, I show the detailed derivation of (1). First, I assume that the recovery rate for the coupon upon default is the same as that of the principal. Formally, I assume

$$\frac{C_{i,t}^f}{C_{i,t}} = \exp(l_{i,t}). \quad (14)$$

Furthermore, I make the technical assumption that after a default occurs, the investor buys the Treasury bond with the coupon rate equal to the original coupon rate, C_i , and short the same bond so that the credit spreads and excess returns are always zero.

I log-linearize returns on corporate bond i such that

$$r_{i,t+1} \approx \rho \delta_{i,t+1} - \delta_{i,t} + \Delta c_{i,t+1} + \text{const}, \quad (15)$$

where $\delta_{i,t} \equiv \log P_{i,t}/C_{i,t}$ and $\Delta c_{i,t+1} \equiv \log C_{i,t+1}/C_{i,t}$.

Similarly, I log-linearize returns on the matching Treasury bonds using the same expansion point, ρ :

$$r_{i,t+1}^f \approx \rho \delta_{i,t+1}^f - \delta_{i,t}^f + \Delta c_{i,t+1}^f + \text{const}, \quad (16)$$

where $\delta_{i,t}^f \equiv \log P_{i,t}^f/C_{i,t}^f$ and $\Delta c_{i,t+1}^f \equiv \log C_{i,t+1}^f/C_{i,t}^f$.

Subtracting (16) from (15) yields

$$r_{i,t+1} - r_{i,t+1}^f \approx -\rho \left(\delta_{i,t+1}^f - \delta_{i,t+1} \right) + \left(\delta_{i,t}^f - \delta_{i,t} \right) - \left(\Delta c_{i,t+1}^f - \Delta c_{i,t+1} \right) + \text{const}. \quad (17)$$

The second term of (17) can be written as

$$\begin{aligned}
\delta_{i,t}^f - \delta_{i,t} &= \log \left(\frac{P_{i,t}^f C_{i,t}}{P_{i,t} C_{i,t}^f} \right), \\
&= \begin{cases} \log \left(\frac{P_{i,t}^f}{P_{i,t}} \right) & \text{if } t \neq t_D \\ 0 & \text{if } t = t_D \end{cases}, \\
&= s_{i,t}.
\end{aligned} \tag{18}$$

In the second equality, I use the fact that the matching Treasury bond has the same coupon rate as the corporate bond, as well as the definition of $l_{i,t}$ in (3) and the assumption in (14).

The last term of (17) is

$$\Delta c_{i,t+1}^f - \Delta c_{i,t+1} = \log \left(\frac{C_{i,t+1}^f C_{i,t}}{C_{i,t+1} C_{i,t}^f} \right).$$

This term can be thought of separately for the three cases: (i) When $t \neq t_D$ and $t+1 \neq t_D$, we have $C_{i,t+1}^f/C_{i,t+1} = C_{i,t}^f/C_{i,t} = 1$ as the matching Treasury bond has the same coupon rate. (ii) When $t \neq t_D$ and $t+1 = t_D$, we have $C_{i,t+1}^f/C_{i,t+1} = \exp(l_{i,t+1})$ by assumption (14), and $C_{i,t}^f/C_{i,t} = 1$. (iii) When $t = t_D$ and $t+1 \neq t_D$, we have $C_{i,t}/C_{i,t}^f = \exp(-l_{i,t})$. However, as I assume that right after the default (time $t+$), the investor buys the bond with the coupon rate equal to C_i , we have $C_{i,t+1}^f = C_{i,t+1} = C_{i,t+}^f = C_{i,t+} = C_i$, so that $\Delta c_{i,t+1}^f - \Delta c_{i,t+1} = \log \left(\frac{C_{i,t+1}^f C_{i,t+}}{C_{i,t+1} C_{i,t+}^f} \right) = 0$. Combining the three cases, we have

$$\Delta c_{i,t+1}^f - \Delta c_{i,t+1} = l_{i,t+1}. \tag{19}$$

Plugging (18) and (19) into (17) leads to the one-period pricing identity in (1).

In the decomposition of the credit spread in (1), there are no terms involving coupon rates, $C_{i,t}$ or $C_{i,t}^f$. Since I work on excess returns rather than returns, the coupons from corporate bonds tend to offset the coupons from the matching Treasury bonds. In addition,

I make the assumption in (14), and thus I completely eliminate the coupon payment from the approximated log excess returns. This feature of the excess returns is convenient as I work on monthly returns. Otherwise, the strong seasonality of coupon payments would make it necessary to use the annual frequency rather than the monthly frequency. Due to the offsetting nature of the excess returns over matching Treasury bonds, I can work on monthly series without adjusting for seasonality.

B Data

B.1 Corporate Bond Database

In this section, I provide a more detailed description of the panel data of corporate bond prices. I obtain monthly price observations of senior unsecured corporate bonds from the following four data sources. First, for the period from 1973 to 1997, I use the Lehman Brothers Fixed Income Database, which provides month-end bid prices. Since Lehman Brothers used these prices to construct the Lehman Brothers bond index while simultaneously trading it, the traders at Lehman Brothers had an incentive to provide correct quotes. Thus, although the prices in the Lehman Brothers Fixed Income Database are quote-based, they are considered reliable.

In the Lehman Brothers Fixed Income Database, some observations are dealers' quotes while others are matrix prices. Matrix prices are set using algorithms based on the quoted prices of other bonds with similar characteristics. Though matrix prices are less reliable than actual dealer quotes (Warga and Welch (1993)), I choose to include matrix prices in our main result to maximize the power of the test. However, I also repeat the main exercise below and show that the results are robust to the exclusion of matrix prices.

Second, for the period from 1994 to 2011, I use the Mergent FISD/NAIC Database. This database consists of actual transaction prices reported by insurance companies. Third, for

the period from 2002 to 2011, I use TRACE data, which provides actual transaction prices. TRACE covers more than 99 percent of the OTC activities in U.S. corporate bond markets after 2005. The data from Mergent FISD/NAIC and TRACE are transaction-based data, and therefore the observations are not exactly at the end of months. Thus, I use only the observations that are in the last five days of each month. If there are multiple observations in the last five days, I use the latest one and treat it as a month-end observation. Lastly, I use the DataStream database, which provides month-end price quotes from 1990 to 2011.

TRACE includes some observations from the trades that are eventually cancelled or corrected. I drop all cancelled observations, and use the corrected prices for the trades that are corrected. I also drop all the price observations that include dealer commissions, as the commission is not reflecting the value of the bond, and these prices are not comparable to the prices without commissions.

Since there are some overlaps among the four databases, I prioritize in the following order: the Lehman Brothers Fixed Income Database, TRACE, Mergent FISD/NAIC and DataStream. The number of overlaps is not large relative to the total size of the data set, with the largest overlaps between TRACE and Mergent FISD making up 3.3% of the non-overlapping observations. To check the data consistency, I examine the effect of priority ordering by reversing the priority, and the effect of the price difference on the empirical result below.

To classify the bonds based on credit ratings, I use the ratings of Standard & Poor's when available, and use Moody's ratings when Standard & Poor's rating is not available.

To identify defaults in the data, I use Moody's Default Risk Service, which provides a historical record of bond defaults from 1970 onwards. The same source also provides the secondary-market value of the defaulted bond one month after the incident. If the price observation in the month when a bond defaults is missing in the corporate bond database, I add the Moody's secondary-market price to my data set in order to include all default

observations in the sample.

B.2 Comparing Overlapping Data Sources

Table 6 compares the summary statistics of the monthly returns of corporate bonds in my sample (Panel A) with the alternative database, which uses the reverse priority (Panel B). Namely, in constructing the alternative database, I prioritize in the following order: DataStream, Mergent FISD/NAIC, TRACE and the Lehman Brothers Fixed Income Database. To see a detailed picture, I tabulate the returns based on credit ratings and time periods. I split the sample into two periods: January 1973 to March 1998 and April 1998 to December 2011. I choose the cutoff of March 1998 because the Lehman Brothers Fixed Income Database is available up to March 1998. As there are more duplicate observations after April 1998, the latter period may show a greater differences between the two priority orders.

Comparing the distribution of bond returns in Panel A with that in Panel B, there is very little difference at any rating category or in any time period. The greatest discrepancy is found in junk bonds from January 1973 to March 1998. The mean for the sample used in this paper is 1.35 percent with the standard deviation of 51.42 percent, while they are 1.20 percent and 35.10 percent in the alternative sample. As the most of the percentiles coincide between the two distributions, the difference comes from the maximum of the distribution.

In the online appendix, I show that the variance decomposition results in Table 2 remain unchanged when I estimate the VAR using the dataset with the reverse priority order. Thus, the results in this paper is not driven by a particular priority order among the databases.

B.3 Construction of Matching Treasury Bonds

In this section, I explain the methodology to construct prices of the matching Treasury bonds. First, I interpolate the Treasury yield curve using cubic splines and construct Treasury zero-

Table 6: Comparing Monthly Corporate Bond Returns (Percent)

Period	Rating	Mean	Median	Std.	Percentile							
					1	5	10	25	75	90	95	99
<i>Panel A: Priority Order = Lehman Brothers, TRACE, Mergent FISD, DataStream</i>												
1973/1	Aaa/Aa	0.75	0.59	7.47	-7.87	-3.88	-2.38	-0.53	1.84	3.68	5.19	10.06
to	A	0.71	0.71	2.59	-6.30	-3.55	-2.26	-0.43	1.85	3.50	4.89	7.92
1998/3	Baa	0.82	0.77	2.64	-5.99	-3.46	-2.15	-0.33	1.96	3.68	5.07	8.15
	HY	1.35	0.95	51.42	-11.82	-4.76	-2.89	-0.21	2.33	4.92	6.90	13.38
	Subtotal	0.88	0.76	23.64	-7.76	-3.86	-2.37	-0.39	1.97	3.87	5.47	9.89
1998/4	Aaa/Aa	0.57	0.59	2.26	-6.24	-2.71	-1.49	-0.06	1.11	2.62	3.89	7.88
to	A	0.63	0.60	2.73	-7.06	-2.92	-1.67	-0.17	1.34	2.98	4.38	8.98
2011/12	Baa	0.66	0.59	14.71	-9.22	-3.26	-1.79	-0.24	1.49	3.18	4.73	10.12
	HY	0.79	0.69	9.04	-14.41	-3.95	-1.70	0.39	1.19	3.29	5.59	15.90
	Subtotal	0.71	0.64	10.33	-10.43	-3.38	-1.71	-0.01	1.33	3.15	4.91	12.04
<i>Panel B: Priority Order = DataStream, Mergent FISD, TRACE, Lehman Brothers</i>												
1973/1	Aaa/Aa	0.74	0.59	7.45	-7.87	-3.87	-2.38	-0.53	1.84	3.67	5.18	10.06
to	A	0.71	0.71	2.59	-6.31	-3.54	-2.25	-0.42	1.84	3.49	4.88	7.93
1998/3	Baa	0.82	0.78	2.64	-6.01	-3.45	-2.13	-0.32	1.95	3.66	5.05	8.16
	HY	1.20	0.95	35.10	-11.82	-4.75	-2.85	-0.21	2.33	4.89	6.89	13.43
	Subtotal	0.85	0.76	16.40	-7.78	-3.85	-2.35	-0.39	1.97	3.85	5.46	9.89
1998/4	Aaa/Aa	0.57	0.59	2.33	-6.61	-2.71	-1.45	-0.03	1.08	2.56	3.86	8.20
to	A	0.68	0.59	16.29	-7.66	-2.84	-1.60	-0.11	1.29	2.89	4.32	9.49
2011/12	Baa	0.72	0.59	22.11	-9.28	-3.12	-1.66	-0.17	1.44	3.07	4.57	9.99
	HY	0.77	0.69	5.29	-14.18	-3.79	-1.57	0.43	1.15	3.19	5.45	15.73
	Subtotal	0.73	0.64	15.25	-10.49	-3.26	-1.60	0.04	1.28	3.05	4.79	12.09

The top panel reports the summary statistics of the (gross) corporate bond returns used in the paper. The bottom panel reports the summary statistics of the data where the priority across the database is reversed (DataStream, Mergent FISD, TRACE, Lehman Brothers). HY is high yield bonds that are rated Ba or below.

coupon curves by bootstrapping. At each month and for each corporate bond in the data set, I construct the future cash flow schedule for the coupon and principal payments. Then I multiply each cash flow by the zero-coupon Treasury bond price with the corresponding time to maturity. I add all of the discounted cash flows to obtain the synthetic Treasury bond price that matches the corporate bond. I do this process for all corporate bonds at each month to obtain the panel data of matching Treasury bond prices. With this method, the credit spread measure is, in principle, unaffected by changes in the Treasury yield curve.

C VAR Coefficients for the Joint Decomposition of Bonds and Stocks, and for the Aggregate Portfolios

In this section, I show the coefficient estimates for the VAR used in Section 4 and 5. In both cases, the dynamics are given by (7). For the joint variance decomposition for bonds and stocks, reported in Section 4, I estimate the VAR using the state vector in (12). The estimated coefficients and their standard errors are presented in Panel A of Table 7. For the variance decomposition including the aggregate variables, reported in Section 5, the state vector in (13) applies. The estimated coefficients and their standard errors are presented in Panel B.

Table 7: Estimated VAR Coefficients: Monthly from 1973 to 2011

Panel A: VAR estimates for bond and stock decomposition, $A \times 100$										
	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t}d_{i,t}^{Baa}$	$\tilde{s}_{i,t}d_{i,t}^{Ba}$	$\tilde{s}_{i,t}d_{i,t}^{B-}$	$\tau\tilde{PD}_{i,t}$	$\tilde{r}_{i,t}^{eq}$	$\tilde{bm}_{i,t}$	R^2	
$\tilde{r}_{i,t+1}^e$	-6.19 (1.86)	2.93 (0.66)	-0.30 (0.31)	0.38 (0.61)	-2.23 (0.69)	-0.05 (0.21)	3.79 (0.37)	-0.07 (0.02)	0.02	
$\tilde{s}_{i,t+1}$	10.11 (2.43)	97.74 (0.68)	0.31 (0.32)	-0.31 (0.61)	-1.53 (0.93)	0.13 (0.22)	-2.72 (0.57)	0.04 (0.02)	0.91	
$\tilde{s}_{i,t+1}d_{i,t+1}^{Baa}$	4.25 (0.48)	0.91 (0.12)	95.11 (0.82)	0.14 (0.26)	-0.72 (0.15)	-0.08 (0.09)	-0.55 (0.07)	0.00 (0.01)	0.90	
$\tilde{s}_{i,t+1}d_{i,t+1}^{Ba}$	2.19 (0.48)	0.16 (0.07)	1.33 (0.23)	92.53 (0.95)	0.09 (0.09)	-0.11 (0.06)	-0.40 (0.09)	0.01 (0.01)	0.85	
$\tilde{s}_{i,t+1}d_{i,t+1}^{B-}$	-1.20 (1.83)	-0.17 (0.09)	0.21 (0.09)	3.63 (0.58)	95.47 (1.02)	0.50 (0.13)	-1.17 (0.52)	0.00 (0.01)	0.87	
$\tau\tilde{PD}_{i,t+1}$	-3.07 (1.22)	1.51 (0.41)	-0.23 (0.46)	1.43 (0.63)	3.15 (0.90)	96.59 (1.06)	-2.46 (0.72)	0.15 (0.05)	0.93	
$\tilde{r}_{i,t+1}^{eq}$	11.00 (2.57)	-0.13 (0.58)	0.83 (0.63)	-0.57 (1.07)	-4.88 (1.11)	-0.86 (0.92)	0.60 (1.28)	0.07 (0.10)	0.00	
$\tilde{bm}_{i,t+1}$	-12.45 (2.81)	2.23 (0.98)	-2.53 (1.03)	0.81 (1.21)	2.81 (1.33)	-1.76 (1.53)	-1.23 (1.38)	98.58 (0.32)	0.97	
Panel B: VAR estimates with the aggregate variables, $A \times 100$										
	$r_{i,t}^e$	$s_{i,t}$	$s_{i,t}d_{i,t}^{Baa}$	$s_{i,t}d_{i,t}^{Ba}$	$s_{i,t}d_{i,t}^{B-}$	$\tau PD_{i,t}$	$r_{EW,t}^e$	$s_{EW,t}$	$\tau PD_{EW,t}$	R^2
$r_{i,t+1}^e$	-4.95 (1.99)	4.64 (1.04)	-1.60 (0.49)	-2.29 (0.82)	-3.76 (0.90)	-0.21 (0.26)	-6.84 (7.26)	0.63 (1.92)	3.93 (2.06)	0.02
$s_{i,t+1}$	12.18 (2.75)	95.16 (1.06)	1.77 (0.50)	2.49 (0.83)	1.17 (0.99)	0.1 (0.23)	-0.48 (7.43)	-0.24 (1.93)	-3.99 (2.06)	0.90
$s_{i,t}d_{i,t}^{Baa}$	5.25 (0.55)	0.73 (0.25)	95.68 (0.82)	-0.17 (0.25)	-0.73 (0.24)	-0.02 (0.12)	-2.85 (2.36)	0.53 (0.62)	-1.67 (0.76)	0.92
$s_{i,t}d_{i,t}^{Ba}$	1.98 (0.63)	-0.09 (0.13)	1.25 (0.15)	95.65 (0.84)	0.41 (0.13)	-0.17 (0.08)	-1.62 (1.34)	0.27 (0.32)	-0.72 (0.32)	0.90
$s_{i,t}d_{i,t}^{B-}$	-1.36 (2.34)	-0.89 (0.24)	0.43 (0.11)	2.68 (0.30)	96.74 (0.81)	0.48 (0.16)	-0.21 (2.36)	1.11 (0.39)	-1.55 (0.40)	0.91
$\tau PD_{i,t+1}$	-3.48 (1.29)	3.13 (0.57)	-1.07 (0.50)	-0.05 (0.46)	1.22 (0.73)	96.55 (1.06)	-1.26 (3.58)	2.58 (1.12)	-2.82 (1.72)	0.93
$r_{EW,t+1}^e$	0	0	0	0	0	0	-8.11 (8.34)	2.21 (2.22)	4.21 (2.36)	0.04
$s_{EW,t+1}$	0	0	0	0	0	0	4.48 (8.18)	97.74 (2.05)	-2.13 (2.11)	0.89
$\tau PD_{EW,t+1}$	0	0	0	0	0	0	-6.80 (4.80)	6.07 (1.42)	94.17 (2.51)	0.98

$r_{i,t}^e$ is the log return on the corporate bonds in excess of the matching Treasury bond, $l_{i,t}$ is the credit loss, $s_{i,t}$ is the credit spread of the corporate bonds, $\tau PD_{i,t}$ is the probability of default implied by the Merton model times the bond's duration, $r_{i,t}^{eq}$ is issuer's equity return in excess of T-bill rate and $bm_{i,t}$ is log issuer's equity book-to-market ratio. $d_{i,t}^\theta$ is a dummy variable for the rating θ . The variables with subscript EW are the equal-weighted average across bonds, computed every month.