

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

YOSHIO NOZAWA\*

October 1, 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.

---

\*Yoshio Nozawa is with the Federal Reserve Board. I am grateful for comments and suggestions from professors and fellow students at the University of Chicago. I express particular thanks to John Cochrane, my dissertation committee chair. I also benefited from the comments by Joost Driessen, Simon Gilchrist, Lars Hansen, Don Kim, Ralph Koijen, Arvind Krishnamurthy, Marcelo Ochoa, Kenneth Singleton (Editor), Pietro Veronesi, Bin Wei, Kenji Wada, Vladimir Yankov and participants in workshops at 10th Annual Risk Management Conference, Aoyama Gakuin, Chicago Booth, Chicago Fed, Erasmus Credit Conference, FRB, Hitotsubashi, New York Fed and the University of Tokyo. The views expressed herein are the author's and do not necessarily reflect those of the Board of Governors of the Federal Reserve System. The author does not have any potential conflicts of interest, as identified in the Journal of Finance disclosure policy.

What drives the cross-sectional variation in credit spreads? Credit spreads are higher when the corporate bond issuer faces a higher default risk and when the discount rate for the corporate bond's cash flows 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 number of observations. 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 credit spreads are driven solely by expected default, the volatility ratio for expected credit loss would be one, while the ratio for the risk premium would be zero. In the data, about half the volatility of credit spreads comes from changing expected excess returns.

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 credit spread volatility of high yield bonds.

In contrast to the security-level results, the drivers for the market-wide variation in credit spreads are mostly risk premiums. The difference arises from 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.

Furthermore, 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 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.

In the 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. With the variance decomposition approach, credit spreads are decomposed into expected credit loss and risk premiums with no unexplained residuals.

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), Duffie and Singleton (1999) and Driessen (2005)), the credit default swap spreads (Longstaff, Mithal and Neis (2005)) or option prices (Culp, Nozawa and Veronesi (2015)). 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 and Williams (2009) and Schwartz (2016) decompose the yield spreads in the sovereign and money markets.

The rest of the article is organized as follows: Section I shows the decomposition of the credit spread of corporate bonds. I describe the data and show the empirical results in Section II. Section III presents a joint variance decomposition of bonds and stocks. Section IV examines the variance decomposition of the bond market portfolio. Section V provides concluding remarks.

## I. Decomposition of Corporate Bond Credit Spreads

### A. 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 takes a long position on an individual corporate bond  $i$  until it matures or defaults and a short position on a Treasury bond with the same cash flows as the corporate bond. 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 defaulted 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

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

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 (including accrued interest) and coupon rate for such a Treasury bond. Then, the return on the matching Treasury bond is

$$R_{i,t+1}^f = \frac{P_{i,t+1}^f + C_{i,t+1}^f}{P_{i,t}^f}. \quad (2)$$

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

$$r_{i,t+1}^e \equiv \log R_{i,t+1} - \log R_{i,t+1}^f \approx -\rho s_{i,t+1} + s_{i,t} - l_{i,t+1} + \text{const}, \quad (3)$$

where

$$s_{i,t} \equiv \begin{cases} \log \frac{P_{i,t}^f}{P_{i,t}} & \text{if } t < t_D, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

$$l_{i,t} \equiv \begin{cases} \log \frac{P_{i,t}^f}{P_{i,t}} & \text{if } t = t_D, \\ 0 & \text{otherwise,} \end{cases} \quad (5)$$

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 (3) 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 (3).

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,  $r_{i,t+1}^e$  can be approximated using a linear function of  $s_{i,t+1}$ ,  $s_{i,t}$  and  $l_{i,t+1}$  in (3) 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 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.<sup>1</sup> Thus, both

---

<sup>1</sup>The average cross-sectional correlation between the price spreads and the yield spreads in my sample is

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 financial 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,<sup>2</sup> 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 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 (3) 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 sensitive to the assumption of the cash flow reinvestment, a point emphasized in Chen (2009).<sup>3</sup>

---

0.82, while the correlation between the price spreads and the yield spreads times duration is 0.97.

<sup>2</sup>For 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."

<sup>3</sup>Chen (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 is not sensitive

The difference equation (3) approximates log excess returns using the first-order Taylor expansion. In the empirical work below, I set the value of  $\rho$  to be 0.992, which minimizes the approximation error in (3). 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,

$$s_{i,t} \approx \sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e + \sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} + \text{const.} \quad (6)$$

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 (6) holds path-by-path, the approximate equality holds under expectation. Taking the time  $t$  conditional expectation of the both sides of (6), we have

$$s_{i,t} \approx E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e \middle| \mathcal{F}_t \right] + E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} \middle| \mathcal{F}_t \right] + \text{const}, \quad (7)$$

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

Equation (7) 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.

---

to the assumption about cash flow reinvestments.



Let us define the long-run expected credit loss as

$$s_{i,t}^l \equiv E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} l_{i,t+j} \middle| \mathcal{F}_t \right]. \quad (8)$$

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,

$$s_{i,t} = s_{i,t}^l + \text{const}, \quad (9)$$

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 in the data is. The expected credit loss hypothesis also implies that the long-run expected excess returns,

$$s_{i,t}^r \equiv E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} r_{i,t+j}^e \middle| \mathcal{F}_t \right], \quad (10)$$

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.

## II. Empirical Results

### A. 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 for a defaulted bond 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, 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 syn-

thetic Treasury bonds that match the corporate bonds using the *Federal Reserve's* constant-maturity yields data. The methodology is detailed in Appendix B.

## B. Estimation by a VAR

I estimate the conditional expectations in (7) 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,

$$X_{i,t} = \begin{pmatrix} \tilde{r}_{i,t}^e & d_{i,t} \tilde{s}_{i,t} & \tau_{i,t} \tilde{z}_{i,t} \end{pmatrix}', \quad (11)$$

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  $\theta$ ,  $\tau_{i,t}$  is the bond's duration and  $z_{i,t}$  is a vector of state variables other than  $\tilde{r}_{i,t}^e$  and  $\tilde{s}_{i,t}$ .

The dynamics of the state variables is given by

$$X_{i,t+1} = AX_{i,t} + W_{i,t+1}. \quad (12)$$

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 for  $\tilde{s}_{i,t}$  to vary as the bond's credit rating changes over time. 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

$$\tilde{s}_{i,t}^l = E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} \tilde{l}_{i,t+j} \middle| X_{i,t} \right] = e_L G(T_i) X_{i,t}, \quad (13)$$

$$\tilde{s}_{i,t}^r = E \left[ \sum_{j=1}^{T_i-t} \rho^{j-1} \tilde{r}_{i,t+j}^e \middle| X_{i,t} \right] = e_1 G(T_i) X_{i,t}, \quad (14)$$

where  $G(T_i) \equiv A(I - \rho A)^{-1} (I - (\rho A)^{T_i-t})$  and  $e_L = -\rho e_2 + e_2 A^{-1} - e_1$ .<sup>4</sup>

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 (6),

$$e_1 G(T_i) + e_L G(T_i) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \end{pmatrix} \quad (15)$$

---

<sup>4</sup>To obtain (13), I use the one-period identity in (3). 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 (13).

holds. Moreover, the expected credit loss hypothesis implies

$$e_1 G(T_i) = \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \end{pmatrix}, \quad (16)$$

$$e_L G(T_i) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \end{pmatrix}, \quad (17)$$

must hold.

Unlike the forecasting coefficients in (15), the volatility ratios for expected credit loss,  $\sigma(s_{i,t}^l) / \sigma(s_{i,t})$ , and expected excess returns,  $\sigma(s_{i,t}^r) / \sigma(s_{i,t})$ , 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.

### C. Main Results

In this section, I estimate the VAR in (12) 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 with rating dummies,  $d_{i,t}\tilde{s}_{i,t}$ , and the probability of default in the Merton model times duration,  $\tau \tilde{P}D_{i,t}$ <sup>5</sup>. (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 zero. I include probability of default based on the Merton (1974) model, because it is known to forecast default (e.g.,

---

<sup>5</sup>  $\tilde{P}D_{i,t} = \Phi(-d_{2,i,t})$ ,

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  $\sigma_A$  is the asset volatility. I use  $rf$  following Bharath and Shumway (2008). Using  $d_{2,i,t}$  in place of the probability of default,  $\tilde{P}D_{i,t}$ , does not change the results.

Gropp, Lo-Duca 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.<sup>6</sup>

Table I 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.

**[Place Table I about here]**

Panel C presents the estimated VAR coefficients. Excess returns tend to be higher when 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, 2.65 (=2.81-0.16) for Baa bonds, and 3.05 (=2.81+0.24) for Ba bonds. The credit spreads and probability of default are fairly autonomous and are forecastable mostly by their own past values.<sup>7</sup>

**[Place Table II about here]**

Panel A of Table II shows the VAR-implied long-run forecasting coefficients in (13) and (14) for a bond with average maturity,  $\bar{T}$ . Holding everything else constant, when credit

---

<sup>6</sup>In the online appendix, I compare the clustered standard errors with the standard errors from bootstrapping which confirms the reliability of the statistical inference.

<sup>7</sup>In the online appendix, I show that the rating transitions implied by the interaction terms,  $E \left[ d_{i,t+1}^{\theta'} \tilde{s}_{i,t+1} \middle| X_{i,t} \right]$ , satisfies a restriction,  $E \left[ d_{i,t+1}^{\theta'} \middle| X_{i,t} \right] \in [0, 1]$ , in the data.

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

**[Place Figure V about here]**

Despite the low R-squared in the return forecasting regression in Panel C of Table I, the return predictability is economically significant: The standard deviation of expected returns is 0.24% per month (not reported in the table) and 5.25% in the long-run (Panel A, Table

II). 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% per month and the difference between different durations is 0.04%.

Panel B of Table II 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 is 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 (3) 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 II 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 (11). 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 0.73 and 0.63. 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 a series of robustness tests. First, the variance decomposition results are robust to the small sample biases. Yu (2002) points out that it is hard to estimate risk premiums on defaultable bonds due to a small number of defaults that we observe. To address this concern, I take a stand on the data generating process by Jarrow, Lando and Turnbull (1997), and show that the variance decomposition based on simulated data can recover the original parameters. Furthermore, the results are 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 II 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.

### III. 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,

$$bm_{i,t} \approx E \left[ \sum_{j=1}^{\infty} \rho_{eq}^{j-1} r_{i,t+j}^{eq} \middle| \mathcal{F}_t \right] + E \left[ \sum_{j=1}^{\infty} \rho_{eq}^{j-1} (y_{i,t+j} - r f_{t+j}) \middle| \mathcal{F}_t \right], \quad (18)$$

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  $\rho_{eq}$  is set to be 0.967, following Vuolteenaho (2002).

Equation (18) 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,

$$X_{i,t} = \begin{pmatrix} \tilde{r}_{i,t}^e & d_{i,t} \tilde{s}_{i,t} & \tau \tilde{P} D_{i,t} & \tilde{r}_{i,t}^{eq} & \tilde{b}m_{i,t} \end{pmatrix}, \quad (19)$$

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

$$\tilde{b}m_{i,t}^y \equiv E \left[ \sum_{j=1}^{\infty} \rho_{eq}^{j-1} (\tilde{y}_{i,t+j} - r f_{i,t+j}) \middle| X_{i,t} \right] = e_y G_{eq}(\infty) X_{i,t}, \quad (20)$$

$$\tilde{b}m_{i,t}^r \equiv E \left[ \sum_{j=1}^{\infty} \rho_{eq}^{j-1} \tilde{r}_{i,t+j}^{eq} \middle| X_{i,t} \right] = e_7 G_{eq}(\infty) X_{i,t}, \quad (21)$$

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  $\tilde{bm}_{i,t}$ ) 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% of the observations (in bond-months) correspond to “Big” stocks that are larger than the 50th NYSE percentile, 12% is “Small” stocks that are between the 20th and 50th NYSE percentiles, while the fraction for “Micro” stocks that are smaller than the 20th NYSE percentile is only 4%. 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.

**[Place Table III about here]**

Table III reports the VAR-implied long-run expectations<sup>8</sup>. Panels A and B show that including stock variables does not materially 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,  $\frac{\sigma(\tilde{bm}^y)}{\sigma(\tilde{bm})}$ , is 1.02, while the volatility ratio for the risk premium,  $\frac{\sigma(\tilde{bm}^r)}{\sigma(\tilde{bm})}$ , 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’

---

<sup>8</sup>The estimates for the one-period VAR coefficients are available upon request.

profitability,  $\rho(\tilde{s}^l, \tilde{bm}^y)$ , is negative and statistically significant. This estimate seems reasonable, as more profitable firms are less likely to default. In contrast, the correlation in risk premiums between bonds and stocks,  $\rho(\tilde{s}^r, \tilde{bm}^r)$ , is insignificant. If the Merton (1974) model holds and leverage is (cross sectionally) constant, then the correlation in risk premiums should be one. The low risk premium correlation is consistent with the existing literature on the weak relationship between bond and stock prices, such as Collin-Dufresne, Goldstein and Martin (2001). The low correlation potentially reflects the variation in leverage or risk-factors other than firms' asset value. Furthermore, liquidity premiums may affect corporate bonds more than stocks, breaking the connection in risk premiums.

Interestingly, the correlation between the expected credit loss for bonds and the risk premiums for stocks,  $\rho(\tilde{s}^l, \tilde{bm}^r)$ , is statistically significantly negative. The firms closer to default earn 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 III, 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.

## IV. Aggregate Credit Spread Dynamics

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

$$r_{EW,t}^e \equiv \frac{1}{N_t} \sum_{i=1}^{N_t} r_{i,t}^e, \quad (22)$$

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:

$$X_{i,t} = \begin{pmatrix} r_{i,t}^e & d_{i,t}s_{i,t} & \tau PD_{i,t} & r_{EW,t}^e & s_{EW,t} & \tau PD_{EW,t} \end{pmatrix}, \quad (23)$$

which follows the dynamics

$$X_{i,t+1} = A_0 + AX_{i,t} + W_{i,t+1}. \quad (24)$$

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.

Panels A and B of Table IV show that the variance decomposition for individual bonds do not change much after including aggregate variables in the VAR<sup>9</sup>. 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).

**[Place Table IV about here]**

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 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:

$$\text{Diversification Factor} = \frac{\sigma^2(s_{EW,t}^u)}{\bar{\sigma}^2(s_{i,t}^u)} \quad u \in \{l, r\}, \quad (25)$$

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

---

<sup>9</sup>The estimates for the one-period VAR coefficients are available upon request.

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<sup>10</sup>. Table IV 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.

## V. 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 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 dominates 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 market portfolio, and their aggregate volatility is small.

Understanding the information in credit spreads is useful for a dynamic portfolio choice

---

<sup>10</sup>The diversification factor may be greater than one as I use an unbalanced panel data. For illustration, consider the case where there are two bonds, and bond A exists in time 1 and 2, while bond B exists in time 3 and 4. Then we can have the diversification factor more than one when the covariance between bonds A and B is large.

problem, as 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, providing a useful signal for portfolio management.

The variance decomposition for credit spreads is useful in examining the link between the corporate bond market and macro economy. Philippon (2009) shows that credit spreads affect corporate investments. Applying the variance decomposition to credit spreads, we can quantify how much reaction in corporate investment is due to changing risk premiums and expected cash flows. I will leave this analysis for future research.

Another 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] 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.
- [2] 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, 645-703.
- [3] Bharath, Sreedhar T. and Tyler Shumway, 2008, Forecasting default with the Merton distance to default model, *Review of Financial Studies* 21, 1339-1369.
- [4] Bongaerts, Dion, 2010, Overrated credit risk: Three essays on credit risk in turbulent times, *Working Paper*.
- [5] Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157-179.
- [6] Campbell, John Y., Jens Hilscher and Jan Szilagyi, 2008, In search of distress risk, *Journal of Finance* 63, 2899-2939.
- [7] 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.
- [8] Campbell, John Y., and Robert J. Shiller, 1988b, Stock prices, earnings, and expected dividends, *Journal of Finance* 43, 661-676.
- [9] Chen, Hui, 2010, Macroeconomic conditions and the puzzles of credit spreads and capital structure, *Journal of Finance* 65, 2171-2212.
- [10] Chen, Long, 2009, On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics* 92, 128-151.
- [11] 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, 3367-3409.
- [12] Cochrane, John H., 2008, The dog that did not bark: A defense of return predictability, *Review of Financial Studies* 21, 1533-1575.
- [13] Cochrane, John H., 2011, Discount rates, *Journal of Finance* 66, 1047-1108.
- [14] Cochrane, John H. and Monika Piazzesi, Bond Risk Premia, *American Economic Review* 95, 138-160.
- [15] Collin-Dufresne, Pierre, Robert S. Goldstein, 2001, Do credit spreads reflect stationary leverage ratios?, *Journal of Finance* 56, 1929-1957.

- [16] Collin-Dufresne, Pierre, Robert S. Goldstein and J. Spencer Martin, 2001, The determinants of credit spread changes, *Journal of Finance* 56, 2177-2207.
- [17] Crabbe, Leland E., 1991, Callable corporate bonds: A vanishing breed, *FEDS working paper #155*, Board of Governors of the Federal Reserve System.
- [18] Culp, Chris, Yoshio Nozawa and Pietro Veronesi, 2015, Option-based credit spreads, NBER Working Paper 20776.
- [19] Driessen, Joost, 2005, Is default event risk priced in corporate bonds?, *Review of Financial Studies* 18, 165-195.
- [20] Duffee, Gregory R., 1999, Estimating the price of default risk, *Review of Financial Studies* 12, 197-226.
- [21] Duffie, Darrell , Leandro Saita and Ke Wang, 2007, Multi-period corporate default prediction with stochastic covariates, *Journal of Financial Economics* 83, 635-665.
- [22] Duffie, Darrell and Kenneth J. Singleton, 1999, Modeling term structures of defaultable bonds, *Review of Financial Studies* 12, 687-720.
- [23] Elton, Edwin J., Martin J. Gruber, Deepak Agrawal, Christopher Mann, 2001, Explaining the rate spread on corporate bonds, *Journal of Finance* 56, 247-277.
- [24] Fama, Eugene F., 1984, Forward and spot exchange rates, *Journal of Monetary Economics* 14, 319-38.
- [25] Fama, Eugene F. and Robert R. Bliss, 1987, The information in long-maturity forward rates, *American Economic Review* 77, 680-692.
- [26] 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.
- [27] Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations, *American Economic Review* 102, 1692-1720.
- [28] 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.
- [29] Harada, Kimie, Takatoshi Ito and Shuhei Takahashi, 2010, Is the distance to default a good measure in predicting bank failures? Case studies, manuscript.
- [30] 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.
- [31] 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.

- [32] Jarrow, Robert A., David Lando and Stuart M. Turnbull, 1997, A Markov model for the term structure of credit risk spreads, *Review of Financial Studies* 10, 481-523.
- [33] Leland, Hayne E., 1994, Corporate debt value, bond covenants, and optimal capital structure, *Journal of Finance* 49, 1213-1252.
- [34] 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.
- [35] Lustig, Hanno, and Adrien Verdelhan, 2007, The cross-section of foreign currency risk premia and consumption growth risk, *American Economic Review* 97, 89-117.
- [36] 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.
- [37] Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449-470.
- [38] Moody's, 1999, *Historical Default Rates of Corporate Bond Issuers, 1920-1998*.
- [39] Moody's, 2011, *Corporate Default and Recovery Rates, 1920-2010*.
- [40] Philippon, Thomas, 2009, The bond market's Q, *Quarterly Journal of Economics* 124, 1011-1056.
- [41] Schwartz, Krista, 2016, Mind the gap: Disentangling credit and liquidity in risk spreads, manuscript, University of Pennsylvania.
- [42] Taylor, John B. and John C. Williams, 2009, A black swan in the money market, *American Economic Journal, Macroeconomics* 1, 58-83.
- [43] Vuolteenaho, Tuomo, 2002, What drives firm-level stock returns?, *Journal of Finance* 57, 233-264.
- [44] Warga, Arthur and Ivo Welch, 1993, Bondholder losses in leveraged buyouts, *Review of Financial Studies* 6, 959-982.
- [45] Yu, Fan, 2002, Modeling expected return on defaultable bonds, *Journal of Fixed Income* 12, 69-81.

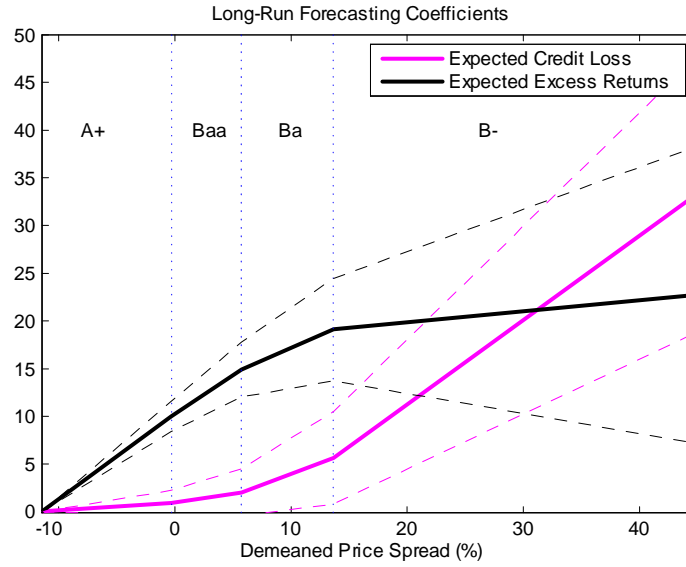


Figure 1: Long-Run Forecasting Coefficients By Credit Ratings

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 II. 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.

# Appendix A. Derivation of the Credit Spread Decomposition

In this appendix, I show the detailed derivation of (3). First, I assume that the recovery rate for the coupon upon default is the same as that of the principal. Formally, I assume

$$\frac{C_{i,t}^f}{C_{i,t}} = \exp(l_{i,t}). \quad (\text{A1})$$

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

$$r_{i,t+1} \approx \rho \delta_{i,t+1} - \delta_{i,t} + \Delta c_{i,t+1} + \text{const}, \quad (\text{A2})$$

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,  $\rho$ :

$$r_{i,t+1}^f \approx \rho \delta_{i,t+1}^f - \delta_{i,t}^f + \Delta c_{i,t+1}^f + \text{const}, \quad (\text{A3})$$

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

Subtracting (A3) from (A2) yields

$$r_{i,t+1} - r_{i,t+1}^f \approx -\rho \left( \delta_{i,t+1}^f - \delta_{i,t+1} \right) + \left( \delta_{i,t}^f - \delta_{i,t} \right) - \left( \Delta c_{i,t+1}^f - \Delta c_{i,t+1} \right) + \text{const}. \quad (\text{A4})$$

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

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%) 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  $\theta$ . The number of observations is 791,864 bond months, and there are 260 default observations. Standard errors, reported in parentheses under each coefficient, are clustered by time.

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
$\tau \tilde{PD}_{i,t}$	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
$\tilde{bm}_{i,t}$	0	65.49	-108.83	-28.54	4.92	32.15	94.00

Panel C: VAR estimates,  $A \times 100$

	$\tilde{r}_{i,t}^e$	$\tilde{s}_{i,t}$	$\tilde{s}_{i,t} d_{i,t}^{Baa}$	$\tilde{s}_{i,t} d_{i,t}^{Ba}$	$\tilde{s}_{i,t} d_{i,t}^{B-}$	$\tau \tilde{PD}_{i,t}$	$R^2$
$\tilde{r}_{i,t+1}^e$	-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
$\tau \tilde{PD}_{i,t+1}$	-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

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

The sample period is monthly from 1973 to 2011. Panel A shows the VAR-implied long-run coefficient for long-run credit loss,  $e_L G(\bar{T})$ , and long-run excess returns,  $e_1 G(\bar{T})$ , where  $G(\bar{T}) = A(I - \rho A)^{-1} (I - (\rho A)^{\bar{T}-t})$  for a bond with the average maturity.  $\sigma(E_t[\cdot])$  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  $\theta$ . The right-hand side variables are defined in the notes to Table I. 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 estimated coefficients for credit loss forecasting regression,  $\tilde{l}_{i,t} = b_l X_{i,t} + \varepsilon_{i,t}^l$ .  $\tilde{l}_{i,t}^l$  is the credit loss implied from the identity (3), so that  $\tilde{l}_{i,t}^l \equiv -\rho s_{i,t} + s_{i,t-1} - r_{i,t}^e$ . 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.

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

	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$
Estimates	0.67 (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
$\tilde{l}_{i,t+1}^l$	-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

	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$
Estimates	0.69 (0.13)	0.54 (0.05)	0.75 (0.03)	0.60 (0.15)	0.10 (0.20)

Panel E: Long VAR

	$\frac{\sigma(\tilde{s}^l)}{\sigma(\tilde{s})}$	$\frac{\sigma(\tilde{s}^r)}{\sigma(\tilde{s})}$	$\varrho(\tilde{s}^l, \tilde{s})$	$\varrho(\tilde{s}^r, \tilde{s})$	$\varrho(\tilde{s}^l, \tilde{s}^r)$
Estimates	0.52 (0.11)	0.59 (0.06)	0.73 (0.05)	0.63 (0.16)	0.24 (0.16)

Table III: Joint Decomposition of Bond and Stock Prices

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

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



Table IV: Decomposition of the Equal-Weighted Market Portfolio

The sample period is monthly from 1973 to 2011. Panel A shows the VAR-implied long-run coefficient for long-run credit loss,  $e_L G(\bar{T})$ , and long-run excess returns,  $e_1 G(\bar{T})$ , where  $G(\bar{T}) = A(I - \rho A)^{-1} (I - (\rho A)^{\bar{T}-t})$  for a bond with the average maturity.  $d_{i,t}^\theta$  is a dummy variable for the rating  $\theta$  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.

Panel A: Long-run regression coefficients,  $e_L G(\bar{T})$  and  $e_1 G(\bar{T})$

	$r_{i,t}^e$	$s_{i,t}$	$s_{i,t} d_{i,t}^{Baa}$	$s_{i,t} d_{i,t}^{Ba}$	$s_{i,t} d_{i,t}^{B-}$
$\sum_{j=1}^{\bar{T}} \rho^{j-1} l_{i,t+j}$	-0.06 (0.02)	0.07 (0.06)	0.15 (0.05)	0.41 (0.13)	0.76 (0.14)
$\sum_{j=1}^{\bar{T}} \rho^{j-1} r_{i,t+j}^e$	0.06 (0.02)	0.93 (0.06)	-0.16 (0.06)	-0.41 (0.13)	-0.76 (0.14)
	$\tau PD_{i,t}$	$r_{EW,t}^e$	$s_{EW,t}$	$\tau PD_{EW,t}$	$\sigma(E_t[\cdot])$
$\sum_{j=1}^{\bar{T}} \rho^{j-1} l_{i,t+j}$	0.10 (0.05)	0.07 (0.02)	-0.20 (0.12)	-0.27 (0.11)	7.91
$\sum_{j=1}^{\bar{T}} \rho^{j-1} r_{i,t+j}^e$	-0.10 (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)	$s_{EW}^l$ 0.08 $s_{EW}^r$ 1.17

The second term of (A4) can be written as

$$\begin{aligned}
\delta_{i,t}^f - \delta_{i,t} &= \log \left( \frac{P_{i,t}^f C_{i,t}}{P_{i,t} C_{i,t}^f} \right), \\
&= \begin{cases} \log \left( \frac{P_{i,t}^f}{P_{i,t}} \right) & \text{if } t \neq t_D \\ 0 & \text{if } t = t_D \end{cases}, \\
&= s_{i,t}.
\end{aligned} \tag{A5}$$

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 (5) and the assumption in (A1).

The last term of (A4) 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 (A1), 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

$$\Delta c_{i,t+1}^f - \Delta c_{i,t+1} = l_{i,t+1}. \tag{A6}$$

Plugging (A5) and (A6) into (A4) leads to the one-period pricing identity in (3).

In the decomposition of the credit spread in (3), 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 (A1), 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.

## Appendix B. Data

### *Appendix A. 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 my main result to maximize the power of the test. However, I also repeat the main exercise in the online appendix 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% 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/NAIC* 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 in the online appendix.

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.

## *Appendix B. Comparing Overlapping Data Sources*

Table V 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 high yield bonds from January 1973 to March 1998. The mean for the sample used in this paper is 1.35% with the standard deviation of 51.42%, while they are 1.20% and 35.10% 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 II 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.

**[Place Table V about here]**

### *Appendix C. 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-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.

Table V: Comparing Monthly Corporate Bond Returns (Percent)

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/NAIC*, *TRACE*, *Lehman Brothers Fixed Income Database*). HY is high yield bonds that are rated Ba or below.

	Mean	Median	Std.	Percentile			
				1	10	90	99
<i>Priority Order: Lehman Brothers, TRACE, Mergent FISD, DataStream</i>							
From 1973/1 to 1998/3							
AAA/AA	0.75	0.59	7.47	-7.87	-2.38	3.68	10.06
A	0.71	0.71	2.59	-6.30	-2.26	3.50	7.92
BBB	0.82	0.77	2.64	-5.99	-2.15	3.68	8.15
HY	1.35	0.95	51.42	-11.82	-2.89	4.92	13.38
From 1998/4 to 2011/12							
AAA/AA	0.57	0.59	2.26	-6.24	-1.49	2.62	7.88
A	0.63	0.60	2.73	-7.06	-1.67	2.98	8.98
BBB	0.66	0.59	14.71	-9.22	-1.79	3.18	10.12
HY	0.79	0.69	9.04	-14.41	-1.70	3.29	15.90
<i>Priority Order: DataStream, Mergent FISD, TRACE, Lehman Brothers</i>							
From 1973/1 to 1998/3							
AAA/AA	0.74	0.59	7.45	-7.87	-2.38	3.67	10.06
A	0.71	0.71	2.59	-6.31	-2.25	3.49	7.93
BBB	0.82	0.78	2.64	-6.01	-2.13	3.66	8.16
HY	1.20	0.95	35.10	-11.82	-2.85	4.89	13.43
From 1998/4 to 2011/12							
AAA/AA	0.57	0.59	2.33	-6.61	-1.45	2.56	8.20
A	0.68	0.59	16.29	-7.66	-1.60	2.89	9.49
BBB	0.72	0.59	22.11	-9.28	-1.66	3.07	9.99
HY	0.77	0.69	5.29	-14.18	-1.57	3.19	15.73