Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

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

Yoshio Nozawa

2014-62

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

Yoshio Nozawa[†]

August 2, 2014

Abstract

I decompose the cross-sectional variation of the credit spreads for corporate bonds into changing expected returns and changing expectation of credit losses with a model-free method. Using a log-linearized pricing identity and a vector autoregression applied to micro-level data from 1973 to 2011, I find that the expected credit loss component and the excess return component each explains about half of the variance of the credit spreads. Unlike the market-level findings in Gilchrist and Zakrajšek (2012), at the firm level, the expected credit loss is volatile and affects the firms' investment decision more than the expected excess returns.

*I am grateful for comments and suggestions from John Cochrane, Simon Gilchrist, Lars Hansen, Don Kim, Arvind Krishnamurthy, Marcelo Ochoa, Bin Wei and participants in workshops at Chicago Booth, Chicago Fed, 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 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 market expectation without relying on a particular model of default. Based on these model-free estimates, I quantify the contributions from the default component and the discount rate component by decomposing the variance of the credit spreads.

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. The log-linear identity that expresses credit spreads as a linear function of log excess returns and credit loss allows me to infer the long-run expected excess returns and the expected credit loss from monthly VARs. Thus, I do not have to take a stand on how firms make their decisions about their capital structure and defaults, or on what factors drive the firm value.

I estimate a VAR involving credit spreads, excess returns, and distance to default of the corporate bonds. Since default occurs infrequently, estimating the expected credit loss and expected returns by running forecasting regressions requires a large dataset. To overcome this issue, I collect corporate bond prices from the Lehman Brothers Fixed Income Database, the Mergent FISD/NAIC Database, TRACE and DataStream, which cover most 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 period.

Based on the estimated VAR, I compute the ratio of volatility of the implied long-run expected credit loss to the volatility of credit spreads and find that the ratio is 0.51. 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. In the data, I find that the estimated volatility ratio of 0.51 is 3.5 standard errors from one. Instead, about half of the volatility of credit spreads comes from changing expected excess returns. The first main empirical finding of the paper is that the expected excess return component is about as volatile as the cash flow component for corporate bonds.

Since the distribution of credit loss is highly skewed, there is a concern about the nonlinearity in the relationship between credit loss and credit spreads. To address the issue of nonlinearity, I estimate the VAR rating-by-rating and allow the slope coefficients to differ across credit ratings. This subsample analysis shows that 97 percent of the variation in credit spreads within investment grade bonds is associated with the variation in risk premium. In contrast, for the subsample of junk bonds, only 31 percent of the credit spread variation corresponds to the discount rate variation, and the rest is due to the expected credit loss. Thus, the relationship between credit loss and credit spreads is nonlinear. However, once I aggregate the rating-based subsample estimates for expected credit loss and compute the volatility ratios using all the bonds, the volatility of expected credit loss is still similar to the volatility of the expected excess returns. Therefore, the first main result remains mostly intact after accounting for nonlinearity.

I run additional VARs with different times to maturity, multiple lags and additional variables such as leverage and equity volatilities, and find that the estimated volatility ratio does not change significantly from the basic VAR specification. These results are robust to the approximation errors, small sample biases, the state tax effect, the effect of call option premium and the inclusion of industry fixed effects. In addition, the return forecasting regressions perform reasonably well out of sample.

The second main result of the paper is the investment forecasting regressions based on the two components of the 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. When the expected credit loss and excess returns are used separately as regressors, both components forecast a decrease in the investment rates. However, when I include both components in a multivariate regression controlling for other determinants of investment, only the expected credit loss predicts investments negatively, and the expected excess returns are statistically insignificant. These results are in stark contrast with the findings of Gilchrist and Zakrajšek (2012), who find that at the aggregate market level, the risk premium component is more important than the default risk component in forecasting macroeconomic activities.

Decomposing credit spreads is important for at least three reasons: First, the variation in the credit spreads is large, and linked to the shocks to the firm-level and macro-level economic activities. As shown by Philippon (2009), Gilchrist, Yankov and Zakrajšek (2009), and Gilchrist, Sim and Zakrajšek (2013), at the aggregate economy level, credit spreads forecast economic fundamentals, such as output, consumption, inflation and investments, above and beyond stock prices do. The variation in credit spreads was even more prominent during the financial crisis in 2008. The difference in yields between Baa and Aaa corporate bonds rose to 260 basis points in October 2008, more than three times as high as the yield difference a year earlier. Due to such a large variation in spreads both over time and across bonds, understanding why the spreads are changing at the individual bond level and market level is important. In this paper, I show that the drivers for the market-wide variation in credit spreads are quite different from the drivers of individual bonds. The difference arises due to 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.

Second, 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.

Third, the examination of the contribution of variation in expected returns on corporate bond prices serves as an out-of-sample test for the excess volatility found in stock prices (e.g., Campbell and Shiller, (1988a and 1998b), Campbell, (1991), Vuolteenaho, (2002), and Cochrane, (2008 and 2011)). In recent studies, Chen (2009) and Chen, Da and Zhao (2013) challenge the previous equity decomposition results by adopting different measures of cash flow news and discount rate news. Thus, an out-of-sample test of the variance decomposition using the securities closest to stocks, namely corporate bonds, contributes to the discussion of stock price volatility. In fact, I find that the variance decomposition results for junk bonds are reasonably close to those for stocks in Vuolteenaho (2002).

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 bond level variation of credit spreads, and decompose the returns on corporate bonds in excess of matching Treasury bond returns to remove shocks to Treasury yield curves. 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 premia. 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.

By examining credit spreads, rather than returns, this article relates to the literature which tries to decompose and explain credit spreads on corporate bonds. Collin-Dufresne, Goldstein and Martin (2001) show that changes in credit spreads cannot be explained by changes in the inputs to the Merton (1974) model, such as leverage and volatility. In addition, numerous papers attempt to explain the credit spread using structural models of debt (e.g., Leland (1994), Collin-Dufresne and Goldstein (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 rely on particular models of default. Instead, by applying Campbell-Shiller (1988a) style decomposition to the credit spread, I estimate the expected credit loss and risk premium components via VARs.

Finally, 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 Piazzezi (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. The difference in interest rates between a home country and a foreign country forecasts returns on the foreign currency, instead of changes in the exchange rate. Beber (2006), McAndrews (2008), Taylor (2009) and Schwartz (2013) decompose the yield spreads in the sovereign and money markets. This article complements the literature by quantifying the variation in risk premia in the corporate bond market.

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 examines how the risk premium and expected credit loss affect firms' investment decisions, and Section 5 provides concluding remarks.

2 Decomposition of the Credit Spreads of Corporate Bonds

2.1 A Simple Example

To illustrate the idea, I start with the simple case of a one-period zero-coupon corporate bond. Suppose there is a one-period corporate bond and a Treasury bond whose face values are normalized to one. At time 0, I observe the log price of the corporate bond, p_0, and the log price of the Treasury bond, p_0^f. Let $s\tau_0$ be the time 0 credit spread, defined as $s\tau_0 \equiv p_0^f - p_0$. At time 1, the corporate bond either matures or defaults. The negative log payoff from the corporate bond at time 1 is given by

$$l_1 = \begin{cases} l > 0 & \text{if defaults,} \\ 0 & \text{otherwise,} \end{cases}$$

while the log payoff from the Treasury bond is always zero. Then, the log returns on the corporate bond and the Treasury bond are

$$\begin{aligned} r_1 &= -l_1 - p_0, \\ r_1^f &= -p_0^f. \end{aligned}$$

Now let us define a log excess return on the corporate bond as

$$r_1^e \equiv r_1 - r_1^f = -l_1 + s\tau_0. \tag{1}$$

Rearranging the terms in (1), I obtain

$$s\tau_0 = r_1^e + l_1. \tag{2}$$

Equation (2) is only a definition of a log excess return and has no economic content. However, this equation provides a useful framework to study the information in the credit spread of a corporate bond. From (2), we can see a simple rule: If the current credit spread is higher, then either the excess return or default loss in the next period must be higher. If $s\tau_0$ varies, either over time or across securities, then $s\tau_0$ must forecast either excess returns or defaults. As (2) holds for any realization of random variables at time 1, the equality also holds under the time 0 conditional expectation:

$$s\tau_0 = E\left[\left.r_1^e\right|\mathcal{F}_0\right] + E\left[\left.l_1\right|\mathcal{F}_0\right], \tag{3}$$

where $\mathcal{F}_0$ is the economic agent's information set. This identity under conditional expectation implies that we can decompose the variation of $s\tau_0$ into the expected excess return component and the expected default component. With equation (3), we can quantify how much of the variation of $s\tau_0$ comes from the risk premium or expected defaults.

I decompose the credit spread using forecasting regressions. If one regresses r_1^e and l_1 on a set of variables in time 0 information set, the fitted values of r_1^e and l_1 are the conditional expectation. This regression-based approach does not rely on a particular model of default. Instead, this methodology uses weighted averages of realized returns and credit losses to estimate the conditional expectation. Combining regression estimates with the identity (3), I can cleanly separate the risk premium component of the credit spread from the expected default component without leaving unexplained residuals. In estimating the expected credit loss, I do not rely on the probability of default estimated by the rating agencies, which is held constant over time. Instead, I allow the expected credit loss and risk premium to vary over time.

The other way to decompose the credit spread is to directly model $E\left[l_1 \mid \mathcal{F}_0\right]$. For example, based on the Merton (1974) model, we can estimate the expected default loss by the model-implied probability of default, multiplied by the loss given default. Once we have $E\left[l_1 \mid \mathcal{F}_0\right]$, we can back out $E\left[r_1^e \mid \mathcal{F}_0\right]$ by $s\tau_0 - E\left[l_1 \mid \mathcal{F}_0\right]$. However, if the model is misspecified, this model-based approach produces a biased decomposition. Thus, in this article, I adopt a model-free approach to decompose credit spreads.

Unlike this simple example, most corporate bonds have a long time to maturity and pay coupons. Furthermore, their expected excess returns and expected defaults may vary over time. Therefore, in the subsection that follows, I develop a more general framework that relates the credit spread of corporate bonds to their excess returns and credit losses under a multi-period setup.

2.2 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 $P^f_{i,t}$ and $C^f_{i,t}$ be the price and coupon rate for such a Treasury bond. Then, the return on the matching Treasury bond is

$$R^f_{i,t+1} = \frac{P^f_{i,t+1} + C^f_{i,t+1}}{P^f_{i,t}}.$$

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. Under this assumption, following Campbell and Shiller (1988a), I log-linearize both $R_{i,t+1}$ and $R^f_{i,t+1}$ 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

$$r^e_{i,t+1} \equiv \log R_{i,t+1} - \log R^f_{i,t+1} \approx -\rho s\tau_{i,t+1} + s\tau_{i,t} - l_{i,t+1} + const, \tag{4}$$

where

$$s\tau_{i,t} \equiv \begin{cases} \log \frac{P^f_{i,t}}{P_{i,t}} & \text{if } t < t_D, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

$$l_{i,t} \equiv \begin{cases} \log \frac{P^f_{i,t}}{P_{i,t}} & \text{if } t = t_D, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

where t_D is the time of default. The variable $s\tau_{i,t}$ measures credit spreads while $l_{i,t}$ measures the credit loss upon default. Equation (4) 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 (4).

The credit spread measure, $s\tau_{i,t}$, is the price spreads rather than the yield spreads. The price spreads have important advantages over the yield spreads: The price spread has a definition based on the simple formula, while yield spreads can only be computed numerically for a bond that pays coupons. As $s\tau_{i,t}$ has a simple analytic form, $s\tau_{i,t}$ can be approximated using a linear function of $s\tau_{i,t+1}, r^e_{i,t+1}$ and $l_{i,t+1}$ in (4) 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, $s\tau_{i,t}$, is closely related to the commonly used yield spread. For coupon bearing bonds, a price change can be approximated by a change in yields multiplied by duration. The 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. 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[1], which is widely used in pricing credit derivatives. However, my measure of credit loss, $l_{i,t}$, is not the loss accrued to the economy due to default, as I do not use the ultimate recovery after the bankruptcy court settlements.

[1] 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 there reasons, we use the trading price of the defaulted instrument as a proxy for the present value of the ultimate recovery."

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. My 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.

In the empirical work below, I set $\rho = \exp(pc) / (1 + \exp(pc))$, where pc is equal to the sample mean of $\log P_{i,t}/C_{i,t}$. Specifically, I use $\rho = 0.993$. The difference equation (4) approximates log excess returns using the first-order Taylor expansion. I show below that the approximation error is small and does not affect my empirical results.

None of the variables on the right-hand side of (4) depend on the coupon payments. Since the corporate bond and the Treasury bond have the same coupon rates, the coupons cancel with each other. As a result, there is no seasonality in these variables, enabling one to use monthly returns for the decomposition. 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 does not rely on the assumption about cash flow reinvestments.

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

$$s\tau_{i,t} \approx \sum_{j=1}^{T-t} \rho^{j-1} r_{i,t+j}^{e} + \sum_{j=1}^{T-t} \rho^{j-1} l_{i,t+j} + const. \tag{7}$$

If the bond defaults at $t_D < T$, 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 with no consequences.

The equation (7) shows that the credit spread of corporate bonds has a discount rate component and a credit loss component. 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 $s\tau_{i,T} = 0$ by construction. As a result, I do not have to impose the condition in which $\rho^j s\tau_{i,t+j}$ tends to zero, as j goes to infinity.

Since (7) holds path-by-path, the approximate equality holds under expectation. Taking the time t conditional expectation of the both sides of (7), we have

$$s\tau_{i,t} \approx E\left[\sum_{j=1}^{T-t} \rho^{j-1} r^e_{i,t+j} \middle| \mathcal{F}_t\right] + E\left[\sum_{j=1}^{T-t} \rho^{j-1} l_{i,t+j} \middle| \mathcal{F}_t\right] + const,$$

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

Let us define the expected credit loss as

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

We can then measure how much the volatility of $s\tau_{i,t}$ corresponds to the volatility of the expected credit loss by the ratio $\sigma\left(s\tau^l_{i,t}\right)/\sigma\left(s\tau_{i,t}\right)$. To evaluate the magnitude of $\sigma\left(s\tau^l_{i,t}\right)/\sigma\left(s\tau_{i,t}\right)$, 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,

$$s\tau_{i,t} = s\tau^l_{i,t} + const,$$

holds.

Under the expected credit loss hypothesis, $\sigma\left(s\tau^{l}_{i,t}\right)/\sigma\left(s\tau_{i,t}\right) = 1$ 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 $s\tau^{r}_{i,t} \equiv E\left[\sum_{j=1}^{T-t} \rho^{j-1} r^{e}_{i,t+j} \middle| \mathcal{F}_t\right]$ is a constant. Under this hypothesis, the long-run excess returns on corporate bonds are not forecastable.

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 premia or irrational expectations, then the excess returns are forecastable using the scaled price.

2.3 Estimation by a VAR

I set up the empirical framework to measure the volatility ratio, 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,

$$X_{i,t} = \begin{pmatrix} \widetilde{r}^{e}_{i,t} & \widetilde{s\tau}_{i,t} & \widetilde{\tau_{i,t} z}_{i,t} \end{pmatrix}',$$

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

The dynamics of the state variables is given by

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

The VAR coefficient matrices A and B are assumed to be 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.

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, $\widetilde{s\tau}_{i,t}$, 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, $\widetilde{s\tau}_{i,t}$ tends to zero as the bond approaches maturity. Thus, we do not have to scale $\widetilde{s\tau}_{i,t}$ with duration. Although I do not scale $\widetilde{r}^e_{i,t}$ with duration, I add another variable $\widetilde{\tau_{i,t} r}^e_{i,t}$ later as a robustness check, and show that adding $\widetilde{\tau_{i,t} r}^e_{i,t}$ in the VAR does not change the results.

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 implied by the VAR is

$$E\left[\sum_{j=1}^{T-t} \rho^{j-1} \widetilde{l}_{t+j} \middle| X_{i,t}\right] = e_L G X_{i,t}, \tag{8}$$

where $G \equiv A(I - \rho A)^{-1}\left(I - (\rho A)^{T-t}\right)$ and $e_L = -\rho e_2 + e_2 A^{-1} - e_1$. To obtain (8), I use the one-period identity in (4). Solving for $\widetilde{l}_{i,t+1}$ and taking the conditional expectation, we have

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

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

In order to diagnose the estimated volatility ratio, it is also useful to consider the implied long-run return forecasting regressions:

$$E\left[\sum_{j=1}^{T-t}\rho^{j-1}\widetilde{r}^{e}_{i,t+j}\middle| X_{i,t}\right]=e_1GX_{i,t}.$$

Then, by identity (7),

$$e_1G+e_LG=\begin{pmatrix}0 & 1 & 0 & \dots & 0\end{pmatrix} \tag{9}$$

holds. Moreover, the expected credit loss hypothesis implies

$$\begin{aligned} e_1G &= \begin{pmatrix}0 & 0 & 0 & \dots & 0\end{pmatrix}, \\ e_LG &= \begin{pmatrix}0 & 1 & 0 & \dots & 0\end{pmatrix}, \end{aligned}$$

must hold. If the estimated volatility ratio is not one, we can examine the long-run regression coefficients to identify the source of the deviation.

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.

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. Thus, removing callable bonds would significantly reduce the length of the sample period, and for this reason I include callable bonds in my sample. As the callable bond price reflects the discount due to the call option value, the yields on these bonds are not exactly comparable to the yields on non callable bonds. 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 Appendix B, 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. These two filters are applied to the prices to buy in. 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. That is, if the same bond jumps up more than 20 percent in one month and comes down more than 20 percent in the following month, I assume that the price observation in the middle is recorded with errors.

To compute excess returns and credit spreads, I need to construct the prices of the synthetic Treasury bonds that match the corporate bonds. To this end, I use the Federal Reserve's constant-maturity yields data. 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.

3.2 Main Results

In this section, I estimate the VAR in the previous section and quantify the contribution of the volatility of expected credit loss to the changes in credit spreads. I start from the simple case in which the state vector includes only $\widetilde{r}^e_{i,t}$, $\widetilde{s\tau}_{i,t}$ and distance to default times duration, $-\widetilde{\tau_{i,t} DD}_{i,t}$. I use $\widetilde{r}^e_{i,t}$ instead of $\widetilde{l}_{i,t}$, as $\widetilde{l}_{i,t}$ in the right-hand side of the regression is mostly zero. I include distance to default because it is known to forecast default (e.g., Gropp, Lo-Duca and Vesala (2006) and Harada, Ito and Takahashi (2010)), and Gilchrist and Zakrajšek (2012) use distance to default to decompose their measure of credit spreads. I use negative distance to default, so a greater value corresponds to a higher probability of default. Appendix C provides the computational details of distance to default.

As I have to take a stand on the maturity of the bonds to use the present-value identity (7), I start by analyzing the sample of five-year bonds. I identify the bonds with the remaining time to maturity of five years, and use their history of data until they drop out of the dataset. For example, if there is a bond whose time to maturity is 20 years at issuance, I use the observations for the bond only after its remaining time to maturity becomes five years. I discard all bonds whose time to maturity at issuance is less than five years. Below, I show that the results are robust to the choice of maturity.

I run pooled OLS regressions using demeaned state variables to forecast credit loss and estimate the VARs. To account for the cross-sectional correlation in error terms, I cluster

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

Variable	Mean	Std.	5%-pct	25%-pct	Median	75%-pct	95%-pct
Panel A: Descriptive Statistics, Basic Data							
$r^e_{i,t}$	0.14	3.37	-2.69	-0.47	0.14	0.74	3.21
$s\tau_{i,t}$	7.73	16.31	0.25	1.26	3.30	8.80	25.03
$\tau_{i,t}DD_{i,t}$	0.20	0.14	0.03	0.10	0.17	0.27	0.47
$l_{i,t}$	0.08	3.82	0.00	0.00	0.00	0.00	0.00
Panel B: Descriptive Statistics, Demeaned Data							
$\widetilde{r}^e_{i,t}$	0	3.25	-2.60	-0.51	-0.04	0.47	3.03
$\widetilde{s\tau}_{i,t}$	0	15.48	-10.96	-4.72	-2.13	1.10	14.29
$\widetilde{\tau_{i,t}DD}_{i,t}$	0	0.13	-0.17	-0.09	-0.02	0.07	0.25
$\widetilde{l}_{i,t}$	0	3.81	-0.43	-0.06	0.00	0.00	0.00

Means, standard deviations and percentiles (5, 25, 50, 75, and 95 percent) are estimated using the monthly panel data of five-year bonds from January 1973 to December 2011. All the variables except $\tau_{i,t}DD_{i,t}$ are shown in percentage. $r^e_{i,t}$ is the log return on the corporate bonds in excess of the matching Treasury bond, $l_{i,t}$ is the credit loss, $s\tau_{i,t}$ is the credit spread of the corporate bonds and $\tau_{i,t}DD_{i,t}$ is the distance to default times the bond's duration. 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. The number of observations is 197,206 bond months.

standard errors by time. Later I compare the clustered standard errors with the standard errors from bootstrapping which confirms the reliability of the statistical inference.

Table 1 shows the summary statistics of the variables used in regressions. The statistics are computed using the panel data of five-year bonds. Panel A shows the raw data before demeaning. The excess returns and distance to default are distributed symmetrically, while the credit spreads and credit loss are right-skewed. Panel B shows the demeaned data, in which the cross-sectional mean is subtracted from each observation. To estimate the VAR below, I use the demeaned data. Demeaning does not significantly reduce the volatility of the variables, while it somewhat reduces the skewness of credit loss.

The first panel of Table 2 shows the estimated credit loss forecasting regressions. In order to see how approximation error affects the regression results, I use both $l_{i,t+1}$ and its log-linear approximation based on (4), $\widetilde{l}'_{i,t+1} \equiv -\rho\widetilde{s\tau}_{i,t+1} + \widetilde{s\tau}_{i,t} - \widetilde{r}^e_{i,t+1}$, for the left-hand side variables.

Table 2 shows that the credit loss measure, $\widetilde{l}_{i,t+1}$, is forecastable based on the credit spread, with a slope coefficient of 2.69. The bond is more likely to default when $\widetilde{s\tau}_{i,t}$ is high (i.e., the price of the corporate bond is relatively lower than the price of the matching Treasury bond). Past excess returns do not forecast the loss next period, while distance to default helps forecast default. When the issuer is closer to default (high $-\widetilde{\tau_{i,t}DD}_{i,t}$), the bond is more likely to default, which is consistent with the Merton (1974) model.

The second panel of Table 2 presents the estimated VAR coefficients. Excess returns tend to be higher when past excess returns are high, credit spreads are high, or the issuer is far from default. Statistically, only the credit spread is significant, with a coefficient of 2.15 and a standard error of 0.98. Though R-squared is low, the return predictability is economically significant, with a standard deviation of expected excess returns of 0.21 percent per month. The variation in expected returns is very 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.

The VAR coefficients for the credit spreads and distance to default show that these two variables are fairly autonomous and are forecastable mostly by their own past values.

Due to the identity (7), lower prices of corporate bonds must correspond to either higher excess returns or higher credit loss. To measure the contribution from these two components, I compute the VAR-implied long-run forecasting coefficients, $e_L G$ and $e_1 G$, shown in the third panel of Table 2. Holding everything else constant, when credit spreads go up by one percent, the expected long-run credit loss only goes up by 0.52 percent. Under the benchmark case of the expected credit loss hypothesis, the slope coefficient on credit spreads must be one. In the data, the estimated long-run credit loss forecasting coefficient is more than three standard errors from one.

Table 2: Estimated VARs, Implied Long-run Regression Coefficients and Volatility Ratios

	Explanatory variable $\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	R^2	Joint significance	$\sigma(E_t[y_{i,t+1}])$
Regression of credit loss on information:						
$\widetilde{l}_{i,t+1}$	-5.63	2.69	1.85	0.03	[0.000]	0.30
	(3.88)	(0.91)	(0.85)			
$\widetilde{l}'_{i,t+1}$	-5.19	2.38	1.86	0.02	[0.018]	0.26
	(3.92)	(0.92)	(0.86)			
VAR estimates: $A \times 100$						
$\widetilde{r}^e_{i,t+1}$	1.05	2.15	-1.79	0.01	[0.000]	0.21
	(3.30)	(0.98)	(1.24)			
$\widetilde{s\tau}_{i,t+1}$	4.17	96.14	-0.07	0.90	[0.000]	
	(5.15)	(1.25)	(1.50)			
$-\tau_{i,t+1}\widetilde{DD}_{i,t+1}$	-0.16	0.05	98.22	0.99	[0.000]	
	(0.03)	(0.01)	(0.29)			
Long-run regression coefficients:						
$\sum \rho^{j-1}\widetilde{l}_{i,t+j}$	-0.03	0.52	0.76			4.76
	(0.02)	(0.16)	(0.34)			
$\sum \rho^{j-1}\widetilde{r}^e_{i,t+j}$	0.03	0.48	-0.76			4.80
	(0.02)	(0.16)	(0.34)			
Implications of VAR estimates:						
	$\sigma(s\tau^l)/\sigma(s\tau)$	$\sigma(s\tau^r)/\sigma(s\tau)$	$corr(s\tau^l, s\tau)$	$corr(s\tau^r, s\tau)$	$corr(s\tau^l, s\tau^r)$	
Estimates	0.51	0.51	0.900	0.900	0.919	
	(0.16)	(0.15)	(0.012)	(0.025)	(0.068)	

The sample period is monthly from 1973 to 2011. $\widetilde{r}^e_{i,t}$ is the log return on the corporate bond i in excess of the matching Treasury bond, $\widetilde{l}_{i,t}$ is the credit loss on bond i, $\widetilde{l}'_{i,t}$ is the credit loss implied from $\widetilde{r}^e_{i,t}, \widetilde{s\tau}_{i,t-1}$ and $\widetilde{s\tau}_{i,t}$ based on (4), $\widetilde{s\tau}_{i,t}$ is the credit spread, $DD_{i,t}$ is the distance to default, and $\tau_{i,t}$ is the bond's duration. The variables $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ are the sum of expected long-run discounted credit loss and the sum of expected long-run discounted excess returns, defined by $s\tau^l_{i,t} = e_L A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and $s\tau^r_{i,t} = e_1 A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. The column $\sigma(E_t[y_{t+1}])$ shows the sample standard deviation of fitted values of the left-hand side variables. Standard errors, reported in parentheses under each coefficient, are clustered by time, and p-values are reported in brackets. The matrix A and the associated standard errors are multiplied by 100.

Since the long-run forecasting coefficients on the credit spreads add up to one, the excess return forecasting coefficient is 1-0.52=0.48, which is three standard errors from zero. Economically, the return forecastability is as large as the credit loss forecastability. When the credit spread increases, expected excess returns go up about as much as expected credit loss does. Though lower corporate bond prices signal higher defaults in the future, the credit loss predictability is not large enough to eliminate the excess return predictability. As a result, rising credit spreads signals both increasing default loss and excess returns in the future. Statistically, the evidence for excess return forecastability is more clear in the implied long-run coefficients than in the one-period coefficients. Since the long-run coefficients are estimated more precisely, they are highly statistically significant, while the one-period coefficients are only marginally significant.

By the identity (9), the slope coefficients of long-run excess returns and long-run credit loss on $\widetilde{r}^e_{i,t}$ and $-\tau_{i,t}\widetilde{DD}_{i,t}$ add up to zero. In order for variables other than the credit spreads to forecast long-run excess returns, these variables have to help credit spreads forecast long-run credit loss. In Table 2, past excess returns do not forecast either long-run excess returns or credit loss, while distance to default has a marginal forecasting power.

The bottom panel shows the ratio of the volatility of expected credit loss to the credit spreads. This volatility ratio answers the question of how much of the volatility of observed credit spreads is driven by changes in expected credit loss. The estimated volatility ratio is 0.51. Since past excess returns and distance to default are economically insignificant forecasters of credit loss, the volatility ratio nearly coincides with the long-run credit loss forecasting coefficient on credit spreads. The ratio of the volatility of long-run expected excess returns to the volatility of credit spreads is 0.51, which is the same as the ratio for the expected credit loss. The volatility ratio for expected excess returns is more than three standard errors from zero, a benchmark value predicted by the expected credit loss hypothesis. Since credit spreads are the only economically significant regressors, long-run expected credit loss and excess returns are highly correlated with credit spreads, with correlation coefficients of 0.90 and 0.90, respectively.

In general, the volatility ratios for expected credit loss and expected excess returns do not have to add up to one. If there are significant forecasters of long-run credit loss other than credit spreads, which in turn forecast long-run excess returns, then the volatility ratios do not add up to one due to the covariance between expected credit loss and expected excess returns. In this VAR specification, the predictive power of past excess returns and distance to default is weak, which makes it possible to cleanly separate the volatility of credit spreads due to expected credit loss from expected excess returns.

The log-linear approximation does not significantly affect the regression results. The regressions coefficients for $\widetilde{l}_{i,t}$ and $\widetilde{l}'_t$ are similar and within one standard error. When I run the VAR using $\widetilde{r}^{e\prime}_{i,t+1} \equiv -\rho \widetilde{s\tau}_{i,t+1} + \widetilde{s\tau}_{i,t} - \widetilde{l}_{i,t+1}$ instead of $\widetilde{r}^{e}_{i,t+1}$, the long-run excess return forecasting coefficient becomes 0.53, instead of 0.48 as in Table 2. Thus, the approximation error does not significantly affect the results, for either the one-period or long-run forecasting coefficients.

There is no particular reason for using five-year bonds in the analysis. To determine whether changing the time to maturity matters, I repeat the VAR estimation using the 3-, 7-, 10-, 15- and 20-year bonds. The results based on the different times to maturity are shown in Table 3. The decomposition result does not change significantly across maturities. The volatilities of expected long-run credit loss and excess returns peak at seven-year bonds. For shorter maturities, the variation of state variables shrinks due to shorter durations, leading to smaller $\sigma\left(s\tau^l\right)$ and $\sigma\left(s\tau^r\right)$. For longer maturities, the sample size becomes limited and slightly biased for high quality bonds, and thus the volatilities for $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ become smaller. The volatility ratio, however, is very stable; the ratio is slightly below 0.5 for expected credit loss and a bit more than 0.5 for expected excess returns.

The last two columns of Table 3 report the VARs for five-year bonds with more lags and more state variables. For more lags, I report the VAR with three lags as an example. The

Table 3: VARs with Various Maturities, Lags and Variables

	VAR(1) with $X_{i,t} = \begin{pmatrix} \widetilde{r}^e_{i,t} & \widetilde{s\tau}_{i,t} & \tau_{i,t}\widetilde{DD}_{i,t} \end{pmatrix}$						More Lags	More Variables
Years	3	5	7	10	15	20	5	5
$\sigma(s\tau^l)$	3.77	4.64	4.99	4.04	3.99	3.71	3.86	4.99
$\sigma(s\tau^l)/\sigma(s\tau)$	0.50	0.50	0.46	0.41	0.43	0.43	0.43	0.53
	(0.17)	(0.15)	(0.10)	(0.10)	(0.12)	(0.12)	(0.13)	(0.19)
$\sigma(s\tau^r)$	4.32	5.27	6.12	5.96	5.34	5.00	5.18	4.40
$\sigma(s\tau^r)/\sigma(s\tau)$	0.57	0.56	0.56	0.60	0.57	0.58	0.57	0.47
	(0.13)	(0.13)	(0.10)	(0.10)	(0.12)	(0.12)	(0.13)	(0.18)

The sample period is monthly from 1973 to 2011. The table shows the estimated volatility ratio based on the VARs, $X_{i,t+1} = AX_{i,t} + BW_{i,t+1}$. The values 3,...,20 mean that the VARs are estimated using the sample of 3,...,20-year bonds. The column 'More Lags' shows the VAR with three lags for five-year bonds, while 'More Variables' shows the estimates for five-year bonds based on the VAR(1) with the state vector $X_{i,t} = \begin{pmatrix} \widetilde{r}^e_{i,t} & \widetilde{s\tau}_{i,t} & \widetilde{\tau_{i,t} r}^e_{i,t} & \tau_{i,t}\widetilde{Lev}_{i,t} & \tau_{i,t}\widetilde{Evol}_{i,t} \end{pmatrix}'$. The long-run expected credit loss is defined by $s\tau^l_{i,t} = e_L A (I - \rho A)^{-1} A \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and the long-run expected returns are defined by $s\tau^r_{i,t} = e_1 A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. Standard errors are clustered by time, and reported in parentheses under each coefficient.

volatility ratios for credit loss and excess returns are 0.43 and 0.57, which are similar to the case of one lag. The 'More Variables' specification extends the state vector to include five variables:

$$X_{i,t} = \begin{pmatrix} \widetilde{r}^e_{i,t} & \widetilde{s\tau}_{i,t} & \widetilde{\tau_{i,t} r}^e_{i,t} & \tau_{i,t}\widetilde{Lev}_{i,t} & \tau_{i,t}\widetilde{Evol}_{i,t} \end{pmatrix}',$$

where $Lev_{i,t}$ is the market leverage defined by the total book value of the issuer's debt divided by the sum of the book value of the debt and the market value of equity, and $Evol_{i,t}$ is the issuer's equity volatility estimated from the daily stock returns over the last year. As past excess returns can affect the risk and expected returns on bonds differently for various maturities, I include excess returns scaled by time as an additional state variable. Since distance to default is essentially a nonlinear function of leverage and equity volatility, I use both variables instead of distance to default to allow for more flexible dependence of default loss on state variables. With more variables, I run the VAR with one lag and find that the results do not change significantly. The volatility ratio for expected long-run credit loss is 0.53, while it is 0.47 for excess returns. The decomposition results are not sensitive to time to maturity of corporate bonds, the lags, or the additional state variables in VARs.

3.3 Subsamples Based on Credit Ratings and Nonlinearity

In this subsection, I estimate the volatility ratio separately for bonds with different credit ratings. Huang and Huang (2012) find that default components implied by structural models explain larger fractions of the level of credit spreads for speculative grade (junk) bonds than for investment grade (IG) bonds. Although I focus on the variation of credit spreads rather than the level of credit spreads, it is still interesting to see how my results differ for IG bonds and junk bonds. In addition, the subsample analysis by credit ratings provides an effective way to address the nonlinearity in the state variables. Nonlinearity can be an issue if extremely high values of credit spreads are more informative about defaults than lower level of spreads. As the credit spread varies across credit ratings, estimating long-run forecasting coefficients separately for each credit rating can handle the potential issues with nonlinearity.

I use the bonds with five years to maturity and split the sample based on the credit ratings when each bond has the time to maturity of five years. As the present value formula in (7) holds for the life of a bond, I keep the bond in the same rating-based subsample even if its credit rating later changes. I use the bonds with the remaining time to maturity of five years, as there are more observations for five-year bonds than bonds with a longer time to maturity.

Using the subset of bonds, I estimate the VAR(1) including $\widetilde{r}^e_{i,t}$, $\widetilde{s\tau}_{i,t}$ and $-\tau_{i,t}\widetilde{DD}_{i,t}$ as state variables. Table 4 shows the standard deviation of $s\tau^l$ and $s\tau^r$ and the volatility ratios for each rating category. The volatility of expected long-run credit loss increases monotonically as credit ratings fall. The volatility is essentially zero for AAA/AA bonds, while it is 20.2 percent for CCC/C bonds. The volatility ratio for expected long-run credit loss also rises from 0.02 to 0.91. In contrast, the volatility of expected long-run excess returns is nearly constant across credit ratings, except for CCC/C bonds. As a result, the volatility ratio for expected excess returns tends to be lower for bonds with a high credit

Table 4: Subsamples Based on Credit Ratings

	AAA/AA	A	BBB	BB	B	CCC/C	IG	Junk
$\sigma(s\tau_l)$	0.06	0.09	1.83	6.32	9.56	20.21	0.23	9.17
$\sigma(s\tau_l)/\sigma(s\tau)$	0.02	0.02	0.32	0.71	0.76	0.91	0.04	0.69
	(0.04)	(0.01)	(0.15)	(0.31)	(0.36)	(0.24)	(0.02)	(0.17)
$\sigma(s\tau_r)$	2.95	3.75	3.98	3.76	3.11	15.99	6.18	4.12
$\sigma(s\tau_r)/\sigma(s\tau)$	0.98	0.98	0.69	0.42	0.25	0.72	0.97	0.31
	(0.02)	(0.01)	(0.14)	(0.21)	(0.36)	(0.17)	(0.02)	(0.17)

Gathering rating-based estimates into one bin

	$\sigma(s\tau^l)$	$\frac{\sigma(s\tau^l)}{\sigma(s\tau)}$	$\sigma(s\tau^r)$	$\frac{\sigma(s\tau^r)}{\sigma(s\tau)}$	$corr(s\tau^l, s\tau^r)$
Estimates	5.11	0.53	4.13	0.43	0.44

The sample period is monthly from 1973 to 2011. The top panel shows the estimated volatility ratios based on the VARs, $X_{i,t+1} = AX_{i,t} + BW_{i,t+1}$, estimated separately for each credit rating. The results are based on five-year bonds and the bonds are sorted into subsamples based on their credit ratings when they are five years to maturity. The long-run expected credit loss is defined by $s\tau^l_{i,t} = e_L GX_{i,t}$ and the long-run expected returns are defined by $s\tau^r_{i,t} = e_1 GX_{i,t}$. IG is the investment grade, which includes AAA/AA, A and BBB, while Junk includes BB, B and CCC/C. Standard errors, reported in parentheses, are clustered by time. The number of observations is 21,797 for AAA/AA, 52,524 for A, 60,960 for BBB, 28,507 for BB, 28,923 for B, 4,495 for CCC/C bonds. The bottom panel collects the separately estimated $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ for AAA/AA to CCC/C into one sample, and compute their summary statistics.

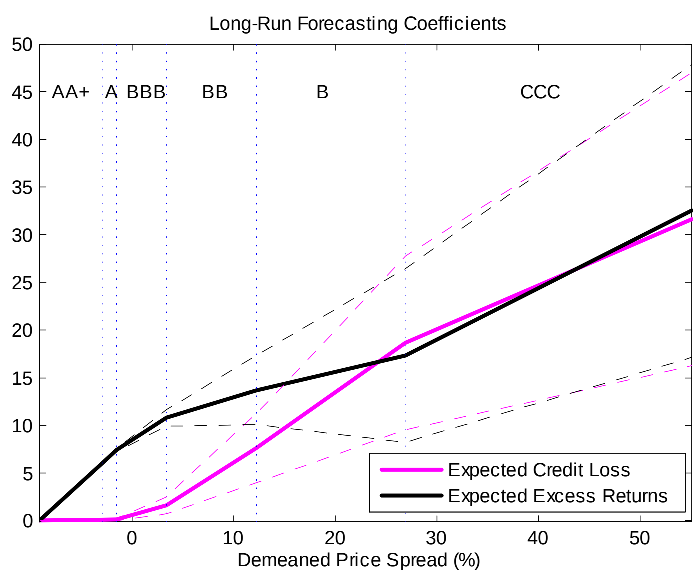
risk.

For the CCC/C bonds, distance to default is volatile and becomes a very strong forecaster of long-run credit loss (and thus long-run excess returns). As a result, both expected long-run credit loss and excess returns are highly volatile compared with credit spreads, and the two volatility ratios do not add up to one.

For IG bonds, 97 percent of the variation in credit spreads corresponds to discount rate news. For Treasury bonds, the ratio $\sigma(s\tau^r)/\sigma(s\tau)$ must be one, as there is no shock to their cash flow. Thus, IG bonds have a volatility ratio similar to Treasury bonds. In contrast, for junk bonds, cash flow news is roughly twice as volatile as the discount rate news. The standard deviation of the expected long-run credit loss for junk bonds is 9.2 percent, which is slightly more than double the volatility for long-run discount rates (4.1 percent). Vuolteenaho (2002) examines the individual stock returns and finds that the ratio is roughly the same: The cash flow news standard deviation is about twice as high as that of the discount rate news. Thus, the behavior of the price of junk bonds is consistent with the previous literature on stocks, while the movement of the IG bond prices is quite similar to that of Treasury bonds.

To examine the effect of the potential nonlinearity, I plot the long-run credit loss and excess return forecasting coefficients on credit spreads, $e_L G e_2'$ and $e_1 G e_2'$ in Figure 1. I determine the range of credit spreads for each credit rating by setting the border in which the two neighboring histograms of credit spreads overlap with each other. Within each range, I draw two lines with the slope equal to $e_L G e_2'$ and $e_1 G e_2'$, while the intercepts are set so that there are no jumps across credit ratings. Figure 1 visualizes how the slope differs across credit ratings and thereby shows the degree of nonlinearity in the long-run VAR.

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, $E_t\left[\sum \rho^{j-1}\widetilde{l}_{i,t+j}\right]$, and excess returns, $E_t\left[\sum \rho^{j-1}\widetilde{r}^{e}_{i,t}\right]$, estimated from the credit rating-based subsamples. Dashed lines denote +/- 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.

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 credit loss starts to steepen, while the line for expected excess returns begins to flatten out. Thus, looking across all ratings, nonlinearity certainly exists, but the ranges of the distribution of the expected credit loss and excess returns are very similar to each other.

In the bottom panel of Table 4, I gather the estimated long-run expected credit loss, $s\tau^l_{i,t} \equiv E_t\left[\sum_j \rho^{j-1}\widetilde{l}_{i,t+j}\right]$, and the estimated long-run expected excess returns, $s\tau^r_{i,t} \equiv E_t\left[\sum_j \rho^{j-1}\widetilde{r}^e_{i,t+j}\right]$, from the rating-based subsamples, and compute volatility across all five-year bonds. This way, I allow the VAR coefficients to be different across the six credit rating categories while computing the volatility ratio using all the rating-based subsamples. Even when I allow the long-run coefficients to differ across ratings, I obtain $\sigma\left(s\tau^l_{i,t}\right) = 5.1$ percent and $\sigma\left(s\tau^r_{i,t}\right) = 4.1$ percent, which are similar to each other. The ratios of these volatilities to the volatility of credit spreads are also similar: 0.53 for credit loss and 0.43 for the expected returns. Therefore, even after allowing the forecasting coefficients to differ across credit ratings and accounting for nonlinearity, the conclusion that the volatility of the cash flow and discount rate shocks is similar still holds in the data.

The subsample analysis based on credit ratings shows that, for an investor who can invest only in IG bonds for institutional reasons, the expected excess returns are the main drivers of the variation in credit spreads. However, for an investor who can invest in bonds with all ranges of credit spreads, to say that the contributions from the expected credit loss component and excess return component are roughly the same is an acceptable description of the overall market.

One result that did change from the simple VAR in Table 2 is the correlation between the expected credit loss and excess returns. When we impose linearity, these two components appear highly correlated, with a correlation coefficient of 0.92. However, this is misleading: Once we account for nonlinearity using the rating-based subsamples as in Table 4, then correlation goes down to 0.44. To correctly estimate the correlation between the expected default and excess returns, it is essential to account for the nonlinearity between the long-run credit loss and credit spreads.

Several readers worry about the “peso problem” in my analysis. If the sample size is too small, then we may not see any defaults for investment grade bonds in the sample, as the probability of default for these bonds is very low. However, because I construct a large panel dataset from the various sources and work on the long-horizon regressions, I still observe defaults in my sample. No AAA bonds jump to default in a month, but since I wait for five years since the bond is rated, the number of defaults in the sample is not zero. In particular, there are two default observations in the AAA/AA subsample and nine default observations in the A subsample. For the A bonds, the volatility of the long-run expected default is statistically significantly different from zero. Thus, the sample size is large enough for variance decomposition, even for investment grade bonds.

3.4 Small Sample Biases and Out-of-Sample Predictability

This section provides several robustness checks. First, the point estimates and standard errors in Table 2 might be affected by a small sample bias, such as the bias pointed out by Stambaugh (1999). The variable, $\widetilde{s\tau}_{i,t}$, is persistent and contemporaneously negatively correlated with excess returns, which may result in potentially spurious return predictability. To address this concern, I run bootstrap simulations, generating 10,000 paths of the state variables under the null that the first row of the estimated matrix A in Table 2 is zero. To bootstrap, I resample months with replacement to generate panel data, allowing heteroskedasticity and cross-sectional correlation in error terms. By running bootstrap simulations instead of Monte Carlo simulations, I allow non-normality in error terms and avoid estimating the variance-covariance matrix of error terms.

Table 5 reports the average point estimates for the key coefficients and test statistics over

Table 5: Bootstrap Simulations

LHV	$\widetilde{r}^e_{i,t+1}$			$\sum \rho^{j-1} \widetilde{r}^e_{i,t+j}$				
RHV	$\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	$\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	$\sigma(s\tau^l)/\sigma(s\tau)$	$\sigma(s\tau^r)/\sigma(s\tau)$
Data:								
Estimates	1.05	2.15	-1.79	0.03	0.48	-0.76	0.51	0.51
Asym. pv	[0.375]	[0.014]	[0.074]	[0.090]	[0.001]	[0.012]	[0.001]	[0.000]
Simulation results:								
Null	0	0	0	0	0	0	1	0
Mean	-0.29	0.02	0.06	0.00	-0.05	-0.05	1.05	0.21
Simulated pv	[0.355]	[0.012]	[0.067]	[0.149]	[0.002]	[0.100]	[0.000]	[0.056]

10,000 simulated paths are generated under the null that $A(1, \cdot) = 0$. Estimates and asymptotic p-values are taken from Table 2. The Mean shows the average of the point estimates over 10,000 paths. Simulated p-values are the fraction of the simulated paths in which the estimates based on the path are below (for $\sigma(s\tau^l)/\sigma(s\tau)$) or above (for other variables) the point estimates from the data.

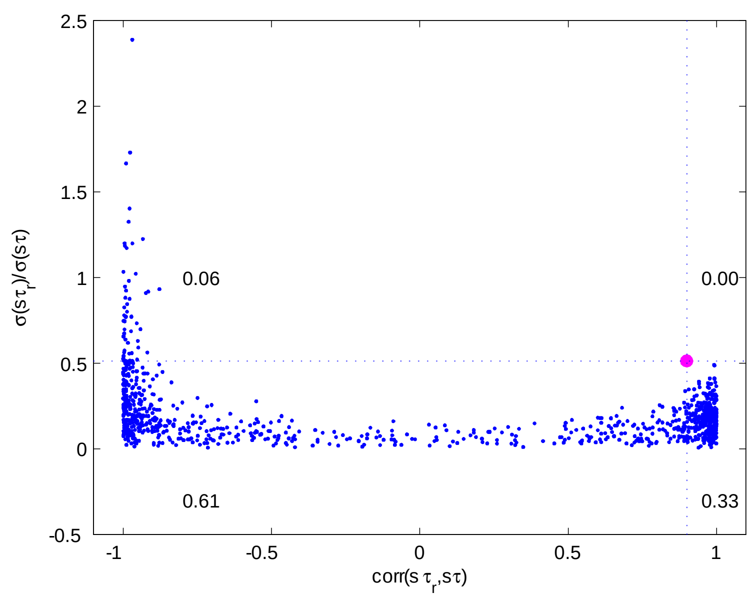
10,000 paths, as well as the fraction of the 10,000 paths that exceed the point estimates in Table 2. The coefficients on $\widetilde{s\tau}_{i,t}$ for the one-period return forecasting regression show a small upward bias of 0.02. However, the p-values based on the simulated probability density function are about the same as the asymptotic p-values. In addition, the long-run return forecasting coefficient on $\widetilde{s\tau}_{i,t}$ is slightly downward biased. Thus, there is little evidence that the contribution of the expected long-run excess return to credit spreads in Table 2 is overestimated.

The volatility ratio for expected long-run credit loss averaged over simulated paths is 1.05, slightly upward biased compared with the null of one. Compared with the point estimate of 0.51, the magnitude of the bias is economically small. In contrast, the volatility ratio for expected long-run excess returns seems to be more upward biased: The average over the simulated paths is 0.21 and the p-value based on the simulated distribution is above 5 percent. However, this upward bias does not contradict the claim that substantial variation in credit spreads corresponds to changing expected returns. The null that expected long-run returns are constant implies that $\sigma(s\tau^r)/\sigma(s\tau) = 0$ and $corr(s\tau^r, s\tau) = 0$. In the data, I find that $\sigma(s\tau^r)/\sigma(s\tau) > 0$ and $corr(s\tau^r, s\tau) > 0$. Thus, to judge the statistical significance of the contribution of the discount rate component, one has to check the probability of observing

$\sigma\left(s\tau^r\right)/\sigma\left(s\tau\right) > 0$ and $corr\left(s\tau^r, s\tau\right) > 0$ at the same time, under the null that long-run returns are not predictable.

Figure 2 plots the volatility ratios and correlations for each simulated path. Most of the paths that give $\sigma\left(s\tau^r\right)/\sigma\left(s\tau\right)$ greater than the sample estimate (0.51) lie in the region of negative correlation, which is counterfactual. On the contrary, to produce a correlation coefficient as large as that in the data, the volatility ratio has to be low. Thus, under the null of constant expected excess returns, the probability that both a large volatility ratio and a high (positive) correlation are drawn is less than one percent. The significance of the expected long-run excess return component in credit spreads is not a spurious result due to small sample bias.

Figure 2: Volatility Ratio and Correlation Based on Bootstrap Simulations



The figure plots the volatility ratio, $\sigma\left(s\tau^r\right)/\sigma\left(s\tau\right)$, and correlation $corr\left(s\tau^r, s\tau\right)$ over the simulated paths. The large dot at (0.9, 0.5) gives the sample estimates. Values are the fraction of 10,000 simulations that fall in the indicated quadrants.

Second, I examine the out-of-sample return predictability. The fact that the expected return variation is a significant driver of the credit spreads suggests that an investor may be able to take advantage of the return predictability in the cross section. In particular, by taking a long position in the bonds with high credit spreads and a short position in the bonds with low credit spreads, the investor should be able to earn higher excess returns, on average. However, as Goyal and Welch (2008) show, many equity return predictors that work in sample do not perform well out of sample. In principle, the out-of-sample test does not validate or invalidate the variance decomposition results. To what extent the price variation comes from changing risk premium is a somewhat different question from the usefulness of return predictability for real-time trading. In fact, Cochrane (2008) shows that even if 100% of the variation in the dividend price ratio of stocks corresponds to the risk premium variation, the out-of-sample return predictability can be rather poor. However, the out-of-sample test can provide a useful diagnostic as to the stability of the forecasting coefficients, and help us detect whether a few outliers drive the entire results.

To examine the out-of-sample return predictability in the cross section, I follow Lewellen (2014) and run the cross-sectional regressions every month

$$\widetilde{r}^e_{i,t+1} = \gamma_t \widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right] + \eta_{i,t+1}, \qquad i = 1, \cdots, N_t,$$

where $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ is the expected excess returns estimated from the same VAR(1) as in Table 2, but using only the information up to month t. If $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ is the true expected return, then the slope coefficient γ_t must be one, while if $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ is just noise and not associated with true expected returns, then we have $\gamma_t = 0$. Of course, even if the VAR(1) is correctly specified, the estimation errors in $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ are likely to bring down γ_t towards zero due to attenuation bias. In this exercise, I examine to what extent the VAR provides a useful measure of the expected returns for real-time trading. In such exercise, investors who try to profit from this trading strategy also suffer from the attenuation biases. Thus, from the investors' perspective, the biased estimate of γ_t is the relevant measure to evaluate the performance of the strategies. In addition, the attenuation bias works against my argument that the forecasting regression performs well out of sample. Thus, I only correct standard errors for the estimation errors in $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$, and leave the coefficient estimates uncorrected.

I run both rolling window forecasting regressions and cumulative forecasting regressions, with the window size of 120 months and 240 months. I repeat the exercise with the full sample and the subsamples for IG bonds and junk bonds. I also estimate VARs separately for AAA/AA, A, BBB, BB, B and CCC/C bonds; collect the estimated $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ for all bonds; and run a single series of monthly cross-sectional regressions to account for potential nonlinearity.

Table 6 shows the average γ_t and R-squared of the out-of-sample forecasting regressions. The estimated expected returns based on the VARs indeed forecast returns out of sample. For all cases except the rolling window regression (120 months) with all bonds, average γ_t is statistically significantly different from zero. For 10 of 16 cases, the average γ_t is statistically indistinguishable from one. The average R-squared ranges from 9% to 18%. Comparing the results for IG bonds and junk bonds, the predictability is pervasive across credit ratings, though the forecasting regression slightly underpredicts the variation in expected returns for the IG bonds and overpredicts the variation for the junk bonds. All told, at least with longer rolling windows (240 months) or cumulative estimation, the forecasting regressions indeed provide a measure of expected returns that forecasts returns out of sample.

In the online appendix[2], I show that the variation in expected excess returns can be explained by exposures to systematic risks. Also, as a further robustness check, I show in the online appendix that the variance decomposition results in this article 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, in the online appendix, I show that the main results in Table 2 are robust, even if I include industry fixed effects in estimating the VAR to account for the difference in credit spreads across industries.

[2] The online appendix can be found at: https://sites.google.com/site/yoshiofinancialeconomics/home/research

Table 6: Return Forecasting Regressions: Out-Of-Sample Performance

Window size	120 months			240 months		
Estimates	γ_t	$s.e.(\gamma_t)$	R^2	γ_t	$s.e.(\gamma_t)$	R^2
Rolling window regressions:						
All	0.34	(0.24)	11.17	0.75	(0.38)	12.84
IG	1.14	(0.13)	15.80	1.18	(0.11)	17.71
Junk	0.70	(0.16)	12.43	0.75	(0.22)	9.94
Gathering AAA to CCC into one group	0.59	(0.08)	9.01	0.68	(0.11)	9.01
Cumulative regressions:						
All	0.90	(0.22)	13.07	0.96	(0.33)	13.29
IG	0.96	(0.09)	16.60	1.11	(0.11)	17.77
Junk	0.86	(0.17)	11.83	0.69	(0.23)	9.92
Gathering AAA to CCC into one group	0.67	(0.08)	9.78	0.69	(0.11)	9.19

Table reports the time-series average of the parameter estimates from the monthly cross-sectional regressions of the form $\widetilde{r}^e_{i,t+1} = \gamma_t \widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right] + \eta_{i,t+1}, i = 1, \cdots, N_t$, where $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ is the expected excess returns estimated using only the information up to t. Rolling window regressions show the estimates in which I estimate $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ using the rolling window of 120 or 240 months. Cumulative regressions show the forecasting results 120 or 240 months after the sample begins, where I use all the data available up to time t to estimate $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$. The row "All" shows the results using all bonds to run a single VAR to estimate $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$, while the rows "IG" and "Junk" show the results using the subsamples of investment grade bonds (AAA/AA, A and BBB) and junk bonds (BB,B and CCC/C). The row "Gathering AAA to CCC into one group" shows the results in which I run a separate VAR for AAA/AA, A, BBB, BB, B and CCC/C to estimate $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$, and gather all the estimates to run a single series of forecasting regressions. R^2 is the (unadjusted) R-squared multiplied by 100. Standard errors, reported in parentheses, are corrected for the estimation errors in $\widehat{E_t}\left[\widetilde{r}^e_{i,t+1}\right]$ using GMM and adjusted for Newey-West 3 lags.

4 Aggregate Credit Spread Dynamics and the Effects on Investment

4.1 VARs with Aggregate Variables

Campbell and Shiller (1988a and 1988b) and Cochrane (2008 and 2011) emphasize the importance of time-varying risk premia 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 just 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 macro variables, and denote them with subscripts EW. For example, the equal-weighted market portfolio returns are computed by

$$r^e_{EW,t} \equiv \frac{1}{N_t} \sum_{i=1}^{N_t} r^e_{i,t},$$

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 macro variables, I run a restricted VAR with a state vector:

$$X_{i,t} = \begin{pmatrix} r^e_{i,t} & s\tau_{i,t} & -\tau_{i,t}DD_{i,t} & r^e_{EW,t} & s\tau_{EW,t} & -\tau DD_{EW,t} \end{pmatrix},$$

which follows the dynamics

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

I restrict the three-by-three entries at the lower left corner of the matrix A to be zero, so that the current individual variables do not forecast the future macro variables. By including the macro variables, I can exploit the cross-sectional variation of individual bonds without demeaning. Thus, based on this VAR with macro 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.

Table 7 shows the estimated VAR and its long-run implications using the five-year bonds. The volatility of expected credit loss is 5.03%, while the volatility of expected excess returns is 6.77%. Thus, these volatility ratios are still comparable to each other. However, the volatility (over time) of the equal-weighted average of expected credit loss is only 0.95%, which is much lower than 4.05% for the equal-weighted average expected excess returns. The reason for the gap between individual bond results and the aggregate market results is the diversification effect. Following Vuolteenaho (2002), I compute the diversification factor:

$$\text{Diversification Factor} = \frac{\sigma^2\left(s\tau^u_{EW,t}\right)}{\bar{\sigma}^2\left(s\tau^u_{i,t}\right)} \qquad u \in \{l, r\},$$

where $\bar{\sigma}^2\left(s\tau^u_{i,t}\right) \equiv \frac{1}{N}\sum_{i=1}^{N}\sigma^2_i\left(s\tau^u_{i,t}\right)$. 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 7 shows that the diversification factor is 0.10 for the expected credit loss while it is 1.19 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 the individual securities.

I repeat the VAR with macro variables using the subsamples based on credit ratings, as shown in Table 8. As before, using the subsamples and allowing nonlinearity do not affect the relative importance of the expected credit loss and excess returns. The overall volatility from the combined estimates of the six subsamples is 6.96% for the expected credit loss and 7.91% for the expected excess returns. The diversification factor is 0.04 for the expected credit loss and 1.24 for the expected excess returns. Only the correlation between expected excess returns and credit loss changes significantly from Table 7. After accounting for nonlinearity, the correlation is close to zero, much lower than 0.599 in Table 7.

4.2 Effects of Credit Spreads on Investment

In this section, I examine how the two components of credit spreads affect firms' investment in the future. Gilchrist and Zakrajšek (2012) decompose credit spreads based on the Merton (1974) model, and find that many macro economic variables are forecastable mainly by the "excess bond premium", or the residuals of credit spreads unexplained by the Merton (1974) model, not by the default risk implied by the model. Gilchrist and Zakrajšek (2012) focus on the aggregate credit spreads, rather than the cross section of individual firms. Thus, it is interesting to see how the different components of the credit spreads forecast economic activities in the cross section.

In this section, I use the decomposition results of the five-year bonds using the credit rating-based subsamples in section 4.1. I use the results based on the rating-based subsamples because the correlation between the expected credit loss and expected excess returns after accounting for nonlinearity is more accurately measured than the full sample results with the linearity assumption. I use the estimates based on the VAR, including macro variables, to contrast the results at the individual firm level and aggregate level.

I take the average of all the bonds issued by a firm to estimate the firm-level expected

Table 7: VAR with Aggregate Variables

	Explanatory variable							Joint significance	
	$r^e_{i,t}$	$s\tau_{i,t}$	$-\tau_{i,t}DD_{i,t}$	$r^e_{EW,t}$	$s\tau_{EW,t}$	$-\tau DD_{EW,t}$	R^2		$\sigma(E_t[y_{t+1}])$
VAR estimates: $A \times 100$									
$r^e_{i,t+1}$	1.63	2.15	-2.39	12.09	0.81	-0.65	0.01	[0.000]	0.33
	(3.35)	(0.95)	(1.11)	(5.51)	(1.06)	(4.87)			
$s\tau_{i,t+1}$	3.48	96.17	0.69	-14.69	0.09	1.34	0.91	[0.000]	
	(5.17)	(1.22)	(1.31)	(5.83)	(1.10)	(4.93)			
$-\tau_{i,t+1}DD_{i,t+1}$	-0.15	0.05	98.16	1.11	-0.46	-2.16	0.99	[0.000]	
	(0.04)	(0.01)	(0.29)	(0.40)	(0.11)	(0.61)			
$r^e_{EW,t+1}$	0	0	0	16.70	3.69	-0.66	0.05	[0.000]	
				(10.35)	(2.52)	(9.46)			
$s\tau_{EW,t+1}$	0	0	0	-22.28	99.95	7.60	0.93	[0.000]	
				(11.00)	(2.61)	(10.03)			
$-\tau DD_{EW,t+1}$	0	0	0	2.37	-0.95	93.67	0.97	[0.000]	
				(0.93)	(0.28)	(1.50)			
Long-run regression coefficients:									
$\sum \rho^{j-1} l_{i,t+j}$	-0.03	0.52	0.85	0.05	-0.47	-0.77			5.03
	(0.02)	(0.16)	(0.34)	(0.07)	(0.36)	(0.70)			
$\sum \rho^{j-1} r^e_{i,t+j}$	0.03	0.48	-0.85	-0.05	0.47	0.77			6.77
	(0.02)	(0.16)	(0.34)	(0.07)	(0.36)	(0.70)			
Implications of VAR estimates:									
	$\frac{\sigma(s\tau^l)}{\sigma(s\tau)}$	$\frac{\sigma(s\tau^r)}{\sigma(s\tau)}$		$corr\left(s\tau^l, s\tau\right)$		$corr\left(s\tau^r, s\tau\right)$		$corr\left(s\tau^l, s\tau^r\right)$	
Estimates	0.48	0.64		0.859		0.925		0.599	
	(0.16)	(0.10)		(0.137)		(0.102)		(0.404)	
Diversification effects:									
		$s\tau^l_{EW}$	$s\tau^r_{EW}$						
Std.		0.95	4.05						
Diversification factor		0.10	1.19						

The sample period is monthly from 1973 to 2011. $\widetilde{r}^e_{i,t}$ is the log return on the corporate bond i in excess of the matching Treasury bond, $\widetilde{l}_{i,t}$ is the credit loss on bond i, $\widetilde{l}'_{i,t}$ is the credit loss implied from $\widetilde{r}^e_{i,t}, \widetilde{s\tau}_{i,t-1}$ and $\widetilde{s\tau}_{i,t}$ based on (4), $\widetilde{s\tau}_{i,t}$ is the credit spread, $DD_{i,t}$ is the distance to default, and $\tau_{i,t}$ is the bond's duration. The variables $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ are the sum of expected long-run discounted credit loss and the sum of expected long-run discounted excess returns, defined by $s\tau^l_{i,t} = e_L A \left(I - \rho A\right)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and $s\tau^r_{i,t} = e_1 A \left(I - \rho A\right)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. The column $\sigma\left(E_t[y_{t+1}]\right)$ shows the sample standard deviation of fitted values of the left-hand side variables. Standard errors, reported in parentheses under each coefficient, are clustered by time, and p-values are reported in brackets. The matrix A and the associated standard errors are multiplied by 100. The variables with subscript EW are the equal-weighted average over bonds, computed every month. The diversification factor is $\sigma^2\left(s\tau^u_{EW,t}\right)/\bar{\sigma}^2\left(s\tau^u_{i,t}\right)$, where $\bar{\sigma}^2\left(s\tau^u_{i,t}\right)$ is the variance of $s\tau^u_{i,t}$ over t averaged across bonds.

Table 8: Subsamples Based on Credit Ratings with Macro Variables

	AAA/AA	A	BBB	BB	B	CCC/C	IG	Junk
$\sigma(s\tau^l)$	0.07	0.14	1.93	6.49	10.08	15.18	1.12	8.92
$\sigma(s\tau^l)/\sigma(s\tau)$	0.02	0.03	0.29	0.60	0.69	0.58	0.20	0.61
	(0.03)	(0.03)	(0.15)	(0.28)	(0.33)	(0.20)	(0.12)	(0.19)
$\sigma(s\tau^r)$	3.34	4.20	5.50	6.35	5.66	14.08	4.75	6.88
$\sigma(s\tau^r)/\sigma(s\tau)$	0.98	0.97	0.83	0.58	0.39	0.54	0.86	0.47
	(0.02)	(0.01)	(0.10)	(0.17)	(0.23)	(0.15)	(0.07)	(0.14)

Gathering rating-based estimates of $s\tau^l$ and $s\tau^r$ into one bin

	$\sigma(\cdot)$	$\sigma(\cdot)/\sigma(s\tau)$	$corr(\cdot, s\tau^r)$		$\sigma(\cdot)$	Div. factor
$s\tau^l$	6.96	0.63	0.01	$s\tau^l_{EW}$	0.72	0.04
$s\tau^r$	7.91	0.71	1	$s\tau^r_{EW}$	3.74	1.24

The sample period is monthly from 1973 to 2011. The top panel shows the estimated volatility ratios based on the VARs with macro variables, $X_{i,t+1} = AX_{i,t} + BW_{i,t+1}$, estimated separately for each credit rating. The results are based on five-year bonds and the bonds are sorted into subsamples based on their credit ratings when they are five years to maturity. The long-run expected credit loss is defined by $s\tau^l_{i,t} = e_L G X_{i,t}$ and the long-run expected returns are defined by $s\tau^r_{i,t} = e_1 G X_{i,t}$. IG is the investment grade, which includes AAA/AA, A and BBB, while Junk includes BB, B and CCC/C. Standard errors, reported in parentheses, are clustered by time. The number of observations is 21,797 for AAA/AA, 52,524 for A, 60,960 for BBB, 28,507 for BB, 28,923 for B, and 4,495 for CCC/C bonds. The bottom panel collects the separately estimated $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ for AAA/AA to CCC/C into one sample, and computes its summary statistics. The variables $s\tau^l_{EW}$ and $s\tau^r_{EW}$ are the equal-weighted average of $s\tau^r_{i,t}$ and $s\tau^r_{i,t}$ over i. Div. factor is the diversification factor defined by $\sigma^2\left(s\tau^u_{EW}\right)/\bar{\sigma}^2\left(s\tau^u_{i,t}\right)$, where $\bar{\sigma}^2\left(s\tau^u_{i,t}\right)$ is the average variance over bond i.

excess returns and credit loss. I forecast the investment rate, measured by the ratio of the capital expenditures this fiscal year to the 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. I also use the book-to-market ratio, profitability, sales-to-capital ratio, idiosyncratic volatility and lagged investment rate of the firm as a control, 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 9 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. Larger expected credit loss and excess returns lead to lower investment next period.

I include all the other control variables in the next two columns. The forecasting power of the expected credit loss increases with more controls, while it decreases for the expected excess returns. When the expected credit loss rises by 1 percent, the investment rate falls by 0.33 percent next year. In contrast, a 1-percent rise in the expected excess returns leads to a 0.09 percent decrease in the investment rate, which is statistically insignificant. The results are similar when I include both the expected credit loss and excess returns in one regression, as shown in the last column. Thus, at the firm level, the expected default component plays a major role in affecting individual firms' investment decisions.

To square with the market-level results in Gilchrist and Zakrajšek (2012), 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 9 shows the estimated coefficients of the forecasting regressions at the market level. When the expected credit loss and excess returns are used separately, but with the lagged investment rate as a control, only the expected excess returns predict a decrease in the investment rate next year. In contrast, the expected credit loss is economically and statistically insignificant. When put together with other control variables, the expected return component becomes an even better predictor of investment, while the expected credit loss component remains insignificant. Figure 3 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 unrelated to the investment rate.

Gilchrist and Zakrajšek (2012) interpret the credit spreads as "a crucial gauge of the degree of strains in the financial system." They argue that "A reduction in the supply of credit—an increase in the excess bond premium—causes a drop in asset prices and a contraction in economic activity through the financial accelerator mechanisms ...." Although it is reasonable to conjecture that much of the risk premium variation in corporate bonds comes from the shocks to financial intermediaries, it is not obvious that firms should react more to the risk premium variation than to the expected cash flow variation. 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. Thus, in an efficient market with no frictions, investment must respond to both types of shocks to asset prices in the same way. However, there are several reasons why the decomposition of shocks to bond prices can show different results in the data.

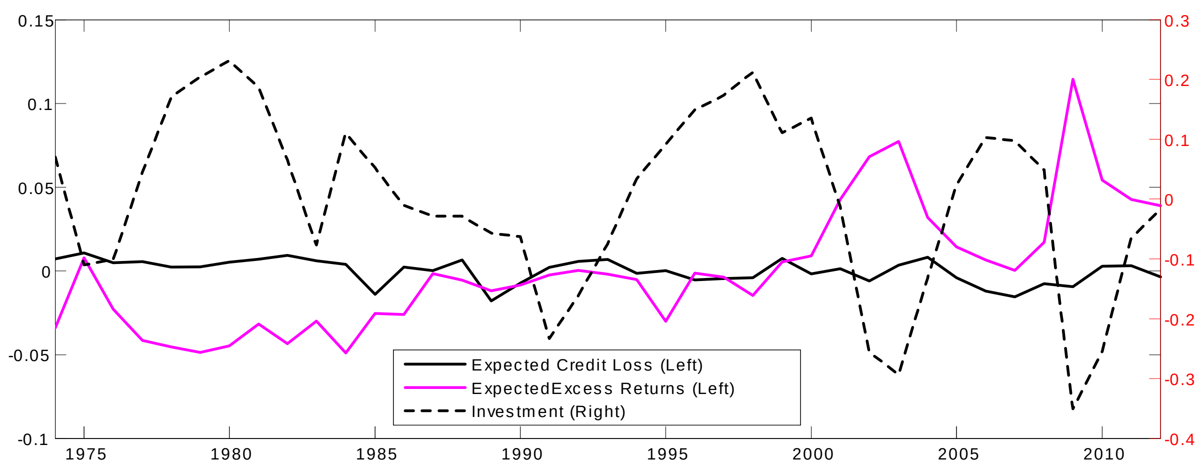
First, the Black-Scholes (1973) and Merton (1974) model implies that the price of a corporate bond is a nonlinear function of the price of the underlying assets. As the option delta changes with the asset values, the price of a corporate bond is an increasing concave function of the underlying assets. This relationship implies that Tobin's average q is a convex function of the credit spreads. Phillipon (2009) confirms this intuition based on a structural model of debt (see his Figure 1). The convexity implies that, for a firm with low spread bonds, a change in the credit spread corresponds to a large change in Tobin's q. In contrast, for a firm with junk bonds, Tobin's q is relatively insensitive to the change in the credit spread. Recall from Table 4 and Table 8 that much of the variation in IG bonds' spread corresponds to the risk premium variation, while the expected default component is more volatile for junk bonds. Taking these pieces of evidence together, we might expect that the risk premium variation, which dominates the credit spread variation for IG bonds, should be more informative about the future investment than the cash flow variation.

Second, 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 value. In the market with irrational investors, the risk premium variation is simply a reflection of time-varying mispricing, which gets corrected over time. Thus, a rational manager should not set the hurdle rate for an investment project based on the short-term fluctuation of market prices and should instead focus on the properties of the expected cash flow from the project. 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. Consistent with this view of the financial market, Greenwood and Hanson (2013) find some evidence for time-varying mispricing in the corporate bond market.

Third, the expected default component can affect firms' investment decisions due to managerial frictions and market segmentations. For example, debt overhang of Myers (1977) suggests that a firm with too much debt may reduce its investment suboptimally, forgoing safe but profitable projects. Debt overhang works even if there is no risk premium, and the debt issued by the firm is fairly priced. If the firm is close to default, the conflict between bond holders and equity holders intensifies, which affects the firm's investment decisions. Thus, a rise in the expected default component can reduce investment through an additional channel due to managerial frictions.

As the three explanations work in opposite directions, which part of the credit spreads better forecasts investment is unclear based on the existing theories. Thus, I have to let the data tell which effects seem to dominate the others. The empirical results above show a striking difference in the information content of the credit spreads at the firm level and at the market level. The shocks to the risk premium are mostly systematic, and these systematic shocks affect firms' collective investment decision. In the aggregate, there is little evidence that the variation in bond mispricing and debt overhang problem affect investment. In contrast, the variation in default risk is the key to understanding the firm-level variation in investment. Thus, my findings are consistent with the interpretation that much of the bond 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.

Figure 3: 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, $s\tau^l_{EW,t}$, and expected excess returns, $s\tau^r_{EW,t}$. The expected credit loss and excess returns are moved forward by one year, so I plot $s\tau^l_{EW,t}$ and $s\tau^r_{EW,t}$ in $t+1$. All variables are demeaned.

Table 9: Investment Forecasting Regressions: Firm-Level Annual Data From 1973 to 2012

	Panel A: Individual Firms Left-hand side variable: $\log I/K_{k,t+12}$					Panel B: Equal-Weighted Market Portfolios Left-hand side variable: $\log I/K_{EW,t+12}$				
$E_t[\sum_j \rho^{j-1} l_{k,t+j}]$	-0.25		-0.33		-0.37	0.13		2.24		-1.37
	(0.11)		(0.12)		(0.15)	(2.34)		(1.97)		(1.29)
$E_t[\sum \rho^{j-1} r^e_{k,t+j}]$		-0.37		-0.09	-0.17		-2.45		-4.09	-4.19
		(0.12)		(0.12)	(0.14)		(0.32)		(0.39)	(0.46)
$\log B/M_{k,t}$			-0.13	-0.13	-0.13			-0.15	-0.13	-0.11
			(0.02)	(0.02)	(0.02)			(0.04)	(0.02)	(0.03)
$\log \Pi/K_{k,t}$			0.03	0.03	0.03			5.30	-0.50	-0.54
			(0.12)	(0.12)	(0.12)			(1.46)	(0.66)	(0.66)
$\log Y/K_{k,t}$			0.06	0.06	0.06			0.49	0.58	0.58
			(0.01)	(0.01)	(0.01)			(0.15)	(0.10)	(0.11)
$\log \sigma^{IV}_{k,t}$			0.02	0.02	0.03			-0.02	0.05	0.06
			(0.02)	(0.02)	(0.02)			(0.03)	(0.03)	(0.03)
$\log I/K_{k,t}$	0.72	0.71	0.64	0.64	0.64	0.55	0.38	0.37	0.09	0.07
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.11)	(0.09)	(0.09)	(0.05)	(0.05)
$\bar{R}^2$	0.519	0.519	0.551	0.551	0.551	0.220	0.590	0.457	0.845	0.843

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 $E_t[\sum_j \rho^{j-1} r^e_{k,t+j}]$ are the long-run expected credit loss and excess returns for the five-year bonds estimated using the credit rating-based subsamples, as in Table 4. 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 $\log B/M_{k,t}$ is the log book-to-market ratio computed following Fama and French (1993), $\log \Pi/K_{k,t}$ 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, $\log \sigma^{IV}_{k,t}$ 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.

5 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 relate the credit spread of corporate bonds to the sum of discounted excess returns and credit loss in the future. Since the relationship among these variables can be log-linearized, I can use VARs to obtain the long-horizon forecasting coefficients and volatility ratios. 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 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 aggregate market results. 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 at the security level is idiosyncratic, the credit loss components are mostly diversified away in the aggregate market, and their aggregate volatility is small.

Consistent with Gilchrist and Zakrajšek (2012), at the market-level, the predictability of investment activities based on the credit spreads comes mostly from the risk premium component. However, at the firm level, the results are the opposite. The expected credit loss component of credit spreads affects individual firm's investment decisions more than the risk premium component does. At the firm-level, the expected credit loss component does not only vary more, but also carry useful information in forecasting future investment activities than at the market level.

One analysis left for the future project 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. In Appendix D, I propose a three-way decomposition of credit spreads in which the spreads are driven by the expected credit loss, excess returns and illiquidity. I show that the model-free approach presented in this paper can be easily extended to account for illiquidity. However, since the illiquidity measures typically require high frequency price observations and/or trading volume data only available from Mergent FISD or TRACE, the sample size will become too small to apply the model-free method. Thus, to decompose the credit spreads taking illiquidity concern into account, we will have to resort to proxies for the default component, such as distance to default or the CDS spreads, or wait until the TRACE data accumulates long enough to cover several credit cycles.

References

[1] Amihud, Yakov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets* 5, 31–56.

[2] 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.

[3] Bao, Jack, Jun Pan and Jiang Wang, 2011, The Illiquidity of Corporate Bonds, *Journal of Finance* 66, 911-946.

[4] 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.

[5] 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.

[6] Bongaerts, Dion, 2010, Overrated Credit Risk: Three Essays on Credit Risk in Turbulent Times, *Working Paper*.

[7] Campbell, John Y., 1991, A Variance Decomposition for Stock Returns, *Economic Journal* 101, 405, 157-179.

[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] Crabbe, Leland E. and Jean Helwege, 1994, Alternative Tests of Agency Theories of Callable Corporate Bonds, *Financial Management* 23, 4, 3-20.

[21] Dick-Nielsen, Jens, Peter Feldhütter and David Lando, 2012, Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis, *Journal of Financial Economics* 103, 471-492.

[22] Driessen, Joost, 2005, Is Default Event Risk Priced in Corporate Bonds?, *Review of Financial Studies* 18, 1, 165-195.

[23] Duffee, Gregory R., 1999, Estimating the Price of Default Risk, *Review of Financial Studies* 12, 197-226.

[24] Duffie, Darrell and Kenneth J. Singleton, 1999, Modeling Term Structures of Defaultable Bonds, *Review of Financial Studies* 12, 687-720.

[25] 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.

[26] Edwards, Amy K., Lawrence E. Harris and Michael S Piwowar, 2007, Corporate Bond Market Transaction Costs and Transparency, *Journal of Finance* 62, 1421-1451.

[27] Fama, Eugene F., 1984, Forward and Spot Exchange Rates, *Journal of Monetary Economics* 14, 319–38.

[28] Fama, Eugene F. and Robert R, Bliss, 1987, The Information in Long-Maturity Forward Rates, *American Economic Review* 77, 4, 680-692.

[29] Feldhütter, Peter, 2012, The Same Bond at Different Prices: Identifying Search Frictions and Selling Pressures, *Review of Financial Studies* 25, 1155-1206.

[30] Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer and Ilya Strebulaev, 2011, Corporate bond default risk: A 150-year perspective, *Journal of Financial Economics* 102, 233-250.

[31] 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.

[32] Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit Spreads and Business Cycle Fluctuations, *American Economic Review* 102, 4, 1692-1720.

[33] Gilchrist, Simon, Jae W. Sim and Egon Zakrajšek, 2013, Uncertainty, Financial Frictions and Irreversible Investment, *Working Paper.*

[34] Goyal, Amit and Ivo Welch, 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, *Review of Financial Studies* 21, 1455-1508.

[35] Greenwood, Robin and Samuel G. Hanson, 2013, Issuer Quality and Corporate Bond Returns, *Review of Financial Studies* 26, 1483–1525.

[36] 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.

[37] Harada, Kimie, Takatoshi Ito and Shuhei Takahashi, 2010, Is the distance to default a good measure in predicting bank failures? Case studies, *Working paper.*

[38] 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.

[39] 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.

[40] Leland, Hayne E., 1994, Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, *Journal of Finance* 49, 4, 1213-1252.

[41] Lewellen, Jonathan W., 2014, The cross section of expected stock returns, *Critical Finance Review*, forthcoming.

[42] 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.

[43] Lin, Hai, Junbo Wang and Chunchi Wu, 2011, Liquidity risk and expected corporate bond returns, *Journal of Financial Economics* 99, 628-650.

[44] Lustig, Hanno, and Adrien Verdelhan, 2007, The Cross-section of Foreign Currency Risk Premia and Consumption Growth Risk, *American Economic Review* 97, 89–117.

[45] 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.

[46] Merton, Robert C., 1974, On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance* 29, 449-470.

[47] Moody's, 1999, Historical Default Rates of Corporate Bond Issuers, 1920-1998.

[48] Moody's, 2011, Corporate Default and Recovery Rates, 1920-2010.

[49] Myers, Stewart C., 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics* 5, 147–175.

[50] Pástor, Ľuboš and Robert F. Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy* 111, 3, 642-685.

[51] Roll, Richard, 1984, A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market, *Journal of Finance* 39, 1127-1139.

[52] Schwartz, Krista, 2013, Mind the Gap: Disentangling Credit and Liquidity in Risk Spreads, *Working Paper.*

[53] Stambaugh, Robert F., 1999, Predictive Regressions, *Journal of Financial Economics* 54, 375-421.

[54] Stein, Jeremy C., 1996, Rational capital budgeting in an irrational world, *Journal of Business* 69, 429-455.

[55] Taylor, John B. and John C. Williams, 2009, A Black Swan in the Money Market, *American Economic Journal, Macroeconomics*, 1: 58-83.

[56] Vuolteenaho, Tuomo, 2002, What Drives Firm-Level Stock Returns?, *Journal of Finance* 57, 1, 233-264.

[57] 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 (4). First, I assume that the recovery rate for the coupon upon default is the same as that of the principal. Formally, I assume

$$\frac{C_{i,t}^f}{C_{i,t}} = \exp\left(l_{i,t}\right). \tag{10}$$

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

$$r_{i,t+1} \approx \rho\delta_{i,t+1} - \delta_{i,t} + \Delta c_{i,t+1} + const, \tag{11}$$

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:

$$r_{i,t+1}^f \approx \rho\delta_{i,t+1}^f - \delta_{i,t}^f + \Delta c_{i,t+1}^f + const, \tag{12}$$

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 (12) from (11) yields

$$r_{i,t+1} - r_{i,t+1}^f \approx -\rho\left(\delta_{i,t+1}^f - \delta_{i,t+1}\right) + \left(\delta_{i,t}^f - \delta_{i,t}\right) - \left(\Delta c_{i,t+1}^f - \Delta c_{i,t+1}\right) + const. \tag{13}$$

The second term of (13) can be written as

$$
\begin{aligned}
\delta^f_{i,t} - \delta_{i,t} &= \log\left(\frac{P^f_{i,t}}{P_{i,t}}\frac{C_{i,t}}{C^f_{i,t}}\right), \\
&= \begin{cases} \log\left(\frac{P^f_{i,t}}{P_{i,t}}\right) & \text{if } t \neq t_D \\ 0 & \text{if } t = t_D \end{cases}, \\
&= s\tau_{i,t}. \qquad (14)
\end{aligned}
$$

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 (6) and the assumption in (10).

The last term of (13) is

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

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 $C^f_{i,t+1}/C_{i,t+1} = C^f_{i,t}/C_{i,t} = 1$ as the matching Treasury bond has the same coupon rate. (ii) When $t \neq t_D$ and $t+1 = t_D$, we have $C^f_{i,t+1}/C_{i,t+1} = \exp(l_{i,t+1})$ by assumption (10), and $C^f_{i,t}/C_{i,t} = 1$. (iii) When $t = t_D$ and $t+1 \neq t_D$, we have $C_{i,t}/C^f_{i,t} = \exp(-l_{i,t})$. 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 $C^f_{i,t+1} = C_{i,t+1} = C^f_{i,t+} = C_{i,t+} = C_i$, so that $\Delta c^f_{i,t+1} - \Delta c_{i,t+1} = \log\left(\frac{C^f_{i,t+1}}{C_{i,t+1}}\frac{C_{i,t+}}{C^f_{i,t+}}\right) = 0$. Combining the three cases, we have

$$
\Delta c^f_{i,t+1} - \Delta c_{i,t+1} = l_{i,t+1}. \qquad (15)
$$

Plugging (14) and (15) into (13) leads to the one-period pricing identity in (4).

In the decomposition of the credit spread in (4), there are no terms involving coupon rates, $C_{i,t}$ or $C^f_{i,t}$. 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 (10), 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 The Effects of Matrix Prices

I repeat the main VAR in Table 2 excluding the observations based on matrix prices. By removing matrix prices, the number of observations decrease to 386,697 bond months from 546,815 bond months in the original dataset. The results without matrix prices are reported in Table 10. The resulting VAR coefficients and volatility ratios are similar to those in Table 2.

B.3 Comparing Overlapping Data Sources

Table 11 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.

Table 10: The Dataset without Matrix Prices

	Explanatory variable $\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	R^2	Joint significance	$\sigma(E_t[y_{t+1}])$
Regression of credit loss on information:						
$\widetilde{l}_{i,t+1}$	-7.62	3.28	2.61	0.03	[0.000]	0.40
	(4.88)	(1.10)	(1.17)			
$\widetilde{l}'_{i,t+1}$	-6.70	2.71	2.32	0.03	[0.009]	0.34
	(4.62)	(0.98)	(1.04)			
VAR estimates: $A \times 100$						
$\widetilde{r}^e_{i,t+1}$	4.93	1.72	-1.91	0.01	[0.000]	0.21
	(3.55)	(1.02)	(1.47)			
$\widetilde{s\tau}_{i,t+1}$	1.79	96.25	-0.42	0.90	[0.000]	
	(5.98)	(1.34)	(1.80)			
$-\tau_{i,t+1}\widetilde{DD}_{i,t+1}$	-0.10	0.04	98.11	0.99	[0.000]	
	(0.03)	(0.01)	(0.35)			
Long-run regression coefficients:						
$\sum \rho^{j-1}\widetilde{l}_{i,t+j}$	-0.06	0.60	0.85			6.01
	(0.03)	(0.18)	(0.41)			
$\sum \rho^{j-1}\widetilde{r}^e_{i,t+j}$	0.06	0.40	-0.85			4.54
	(0.03)	(0.18)	(0.41)			
Implications of VAR estimates:						
	$\sigma(s\tau_l)/\sigma(s\tau)$	$\sigma(s\tau_r)/\sigma(s\tau)$	$corr(s\tau_l, s\tau)$	$corr(s\tau_r, s\tau)$	$corr(s\tau_l, s\tau_r)$	
Estimates	0.58	0.44	0.915	0.903	0.909	
	(0.18)	(0.16)	(0.012)	(0.041)	(0.092)	

The sample period is monthly from 1973 to 2011. $\widetilde{r}^e_{i,t}$ is the log return on the corporate bond i in excess of the matching Treasury bond, $\widetilde{l}_{i,t}$ is the credit loss on bond i, $\widetilde{l}'_{i,t}$ is the credit loss implied from $\widetilde{r}^e_{i,t}, \widetilde{s\tau}_{i,t-1}$ and $\widetilde{s\tau}_{i,t}$ based on (4), $\widetilde{s\tau}_{i,t}$ is the credit spread, $DD_{i,t}$ is the distance to default, and $\tau_{i,t}$ is the bond's duration. The variables $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ are the sum of expected long-run discounted credit loss and the sum of expected long-run discounted excess returns, defined by $s\tau^l_{i,t} = e_L A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and $s\tau^r_{i,t} = e_1 A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. The column $\sigma(E_t[y_{t+1}])$ shows the sample standard deviation of fitted values of the left-hand side variables. Standard errors, reported in parentheses under each coefficient, are clustered by time, and p-values are reported in brackets. The matrix A and the associated standard errors are multiplied by 100.

Table 11: Comparing Monthly Corporate Bond Returns (Percent)

					Percentile							
Period	Rating	Mean	Median	Std.	1	5	10	25	75	90	95	99
Panel A: Priority Order = Lehman Brothers, TRACE, Mergent FISD, DataStream												
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	BBB	0.82	0.77	2.64	-5.99	-3.46	-2.15	-0.33	1.96	3.68	5.07	8.15
	junk	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	BBB	0.66	0.59	14.71	-9.22	-3.26	-1.79	-0.24	1.49	3.18	4.73	10.12
	junk	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
Panel B: Priority Order = DataStream, Mergent FISD, TRACE, Lehman Brothers												
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	BBB	0.82	0.78	2.64	-6.01	-3.45	-2.13	-0.32	1.95	3.66	5.05	8.16
	junk	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	BBB	0.72	0.59	22.11	-9.28	-3.12	-1.66	-0.17	1.44	3.07	4.57	9.99
	junk	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).

To examine how the choices among duplicate data points may affect the final results, I repeat the exercise in Table 2 using an alternative dataset constructed from the reverse priority order. Table 12 reports the estimates of the VAR as well as the test results. The test results are essentially the same as the results in Table 2. Therefore, I conclude that the choice among different datasets does not significantly affect the conclusion of the paper.

B.4 The Effect of Callability

Finally, I show that the main result in Table 2 is robust to the inclusion of the fixed effects of callable bonds. To this end, I demeaned all variables in the state vector $X_{i,t}$ using the cross-sectional averages in each month separately for callable bonds and noncallable bonds. By demeaning separately, callable bonds are allowed to have different means than noncallable bonds. After accounting for callability, I repeat the estimation process in Table 2. The VAR estimated using the separately demeaned data is shown in Table 13. The resulting VAR coefficients and volatility ratios are nearly identical to the estimates in Table 2. Therefore, the difference between callable and noncallable bonds is not driving the main result.

C Computation of Distance to Default

To construct distance to default, I use an implication of the Merton (1974) model.

The value of the assets of a firm, A_t, follows a geometric Brownian motion:

$$\frac{dA_t}{A_t} = \mu dt + \sigma_A dW_t. \tag{16}$$

Let D_t be the book value of the debt of the firm at time t. If the value of the firm's assets is less than the book value of the debt at the maturity date, then it cannot repay the debt and defaults. When in default, the bondholders immediately take over the firm,

Table 12: VARs Based on the Alternative Dataset

	Explanatory variable $\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	R^2	Joint significance	$\sigma(E_t[y_{t+1}])$
Regression of credit loss on information:						
$\widetilde{l}_{i,t+1}$	-5.79	2.74	1.87	0.03	[0.000]	0.30
	(3.89)	(0.93)	(0.87)			
$\widetilde{l}'_{i,t+1}$	-5.41	2.44	1.88	0.02	[0.016]	0.27
	(3.94)	(0.94)	(0.87)			
VAR estimates: $A \times 100$						
$\widetilde{r}^e_{i,t+1}$	2.96	2.25	-1.33	0.01	[0.000]	0.22
	(3.46)	(0.97)	(1.31)			
$\widetilde{s\tau}_{i,t+1}$	2.47	95.98	-0.56	0.90	[0.000]	
	(5.52)	(1.25)	(1.56)			
$-\tau_{i,t+1}\widetilde{DD}_{i,t+1}$	-0.15	0.05	98.22	0.99	[0.000]	
	(0.03)	(0.01)	(0.29)			
Long-run regression coefficients:						
$\sum \rho^{j-1}\widetilde{l}_{I,t+j}$	-0.04	0.51	0.67			4.65
	(0.02)	(0.16)	(0.33)			
$\sum \rho^{j-1}\widetilde{r}^e_{I,t+j}$	0.04	0.49	-0.67			4.86
	(0.02)	(0.16)	(0.33)			
Implications of VAR estimates:						
	$\sigma(s\tau^l)/\sigma(s\tau)$	$\sigma(s\tau^r)/\sigma(s\tau)$	$corr(s\tau^l, s\tau)$	$corr(s\tau^r, s\tau)$	$corr(s\tau^l, s\tau^r)$	
Estimates	0.50	0.52	0.902	0.903	0.936	
	(0.15)	(0.15)	(0.012)	(0.020)	(0.060)	

The sample period is monthly from 1973 to 2011. $\widetilde{r}^e_{i,t}$ is the log return on the corporate bond i in excess of the matching Treasury bond, $\widetilde{l}_{i,t}$ is the credit loss on bond i, $\widetilde{l}'_{i,t}$ is the credit loss implied from $\widetilde{r}^e_{i,t}, \widetilde{s\tau}_{i,t-1}$ and $\widetilde{s\tau}_{i,t}$ based on (4), $\widetilde{s\tau}_{i,t}$ is the credit spread, $DD_{i,t}$ is the distance to default, and $\tau_{i,t}$ is the bond's duration. The variables $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ are the sum of expected long-run discounted credit loss and the sum of expected long-run discounted excess returns, defined by $s\tau^l_{i,t} = e_L A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and $s\tau^r_{i,t} = e_1 A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. The column $\sigma(E_t[y_{t+1}])$ shows the sample standard deviation of fitted values of the left-hand side variables. Standard errors, reported in parentheses under each coefficient, are clustered by time, and p-values are reported in brackets. The matrix A and the associated standard errors are multiplied by 100.

Table 13: Accounting for Call Fixed Effects

Explanatory variable	$\widetilde{r}^e_{i,t}$	$\widetilde{s\tau}_{i,t}$	$-\tau_{i,t}\widetilde{DD}_{i,t}$	R^2	Joint significance	$\sigma(E_t[y_{t+1}])$
Regression of credit loss on information:						
$\widetilde{l}_{i,t+1}$	-5.64	2.76	2.00	0.03	[0.000]	0.30
	(3.90)	(0.93)	(0.91)			
$\widetilde{l'}_{i,t+1}$	-5.19	2.45	2.00	0.02	[0.019]	0.27
	(3.94)	(0.94)	(0.91)			
VAR estimates: $A \times 100$						
$\widetilde{r}^e_{i,t+1}$	0.98	2.18	-1.54	0.01	[0.000]	0.21
	(3.32)	(1.00)	(1.27)			
$\widetilde{s\tau}_{i,t+1}$	4.24	96.04	-0.46	0.90	[0.000]	
	(5.18)	(1.29)	(1.55)			
$-\tau_{i,t+1}\widetilde{DD}_{i,t+1}$	-0.16	0.05	98.22	0.99	[0.000]	
	(0.03)	(0.01)	(0.29)			
Long-run regression coefficients:						
$\sum \rho^{j-1}\widetilde{l}_{i,t+j}$	-0.03	0.52	0.73			4.73
	(0.02)	(0.16)	(0.33)			
$\sum \rho^{j-1}\widetilde{r}^e_{i,t+j}$	0.03	0.48	-0.73			4.71
	(0.02)	(0.16)	(0.33)			
Implications of VAR estimates:						
	$\sigma(s\tau^l)/\sigma(s\tau)$	$\sigma(s\tau^r)/\sigma(s\tau)$	$corr(s\tau^l, s\tau)$	$corr(s\tau^r, s\tau)$	$corr(s\tau^l, s\tau^r)$	
Estimates	0.51	0.51	0.890	0.890	0.923	
	(0.16)	(0.15)	(0.012)	(0.024)	(0.066)	

The sample period is monthly from 1973 to 2011. $\widetilde{r}^e_{i,t}$ is the log return on the corporate bond i in excess of the matching Treasury bond, $\widetilde{l}_{i,t}$ is the credit loss on bond i, $\widetilde{l'}_{i,t}$ is the credit loss implied from $\widetilde{r}^e_{i,t}, \widetilde{s\tau}_{i,t-1}$ and $\widetilde{s\tau}_{i,t}$ based on (4), $\widetilde{s\tau}_{i,t}$ is the credit spread, $DD_{i,t}$ is the distance to default, and $\tau_{i,t}$ is the bond's duration. The variables $s\tau^l_{i,t}$ and $s\tau^r_{i,t}$ are the sum of expected long-run discounted credit loss and the sum of expected long-run discounted excess returns, defined by $s\tau^l_{i,t} = e_L A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$ and $s\tau^r_{i,t} = e_1 A (I - \rho A)^{-1} \left(I - (\rho A)^{T-t}\right) X_{i,t}$. The column $\sigma(E_t[y_{t+1}])$ shows the sample standard deviation of fitted values of the left-hand side variables. Standard errors, reported in parentheses under each coefficient, are clustered by time, and p-values are reported in brackets. The matrix A and the associated standard errors are multiplied by 100. All state variables are demeaned every month using the means separately estimated for callable and non callable bonds.

and the equity holders receive zero. If the assets exceed the debt, then the equity holders receive the difference between A_t and D_t. This way, the market value of equity, S_t, can be considered the price of a call option. The equity value is given by the Black-Scholes formula for a call option:

$$S_t = A_t \Phi(d_1) - D_t \Phi(d_2), \tag{17}$$

where

$$\begin{aligned} d_1 =& \frac{\log(A_t/D_t) + \left(r + \frac{1}{2}\sigma_A^2\right)}{\sigma_A}, \\ d_2 =& \frac{\log(A_t/D_t) + \left(r - \frac{1}{2}\sigma_A^2\right)}{\sigma_A}, \end{aligned}$$

r is a risk-free rate and Φ is a cumulative density function of a standard normal distribution.

We cannot directly observe the market value of the asset, A_t, and its volatility, σ_A. Instead, we can observe the market value of equity, S_t, and its volatility, σ_S. By applying Ito's lemma to the equity and imposing a no-arbitrage condition, we have the risk-neutral dynamics of equity, S_t:

$$dS_t = rS_t dt + \frac{\partial S_t}{\partial A_t} A_t \sigma_A dW_t.$$

By matching the standard deviation of the dynamics, we obtain

$$\begin{aligned} \sigma_S =& \frac{\partial S_t}{\partial A_t} \frac{A_t}{S_t} \sigma_A \\ =& \Phi(d_1) \frac{A_t}{S_t} \sigma_A. \end{aligned} \tag{18}$$

Equations (17) and (18) give a system of two equations with two unknowns: A_t and σ_A.

Since they are nonlinear equations, I solve them numerically using a KNITRO solver. The distance to default is then obtained by

$$-DD_t \equiv -d_2 = -\frac{\log(A_t/D_t) + \left(r - \frac{1}{2}\sigma_A^2\right)}{\sigma_A}.$$

D Decomposition into Three Components, Including Liquidity

The decomposition of the credit spread in (4) can easily be extended to include liquidity. Suppose that the bond market is illiquid and the investor can only buy a bond at the frictionless price times $H_{i,t} \geq 1$. When she sells a bond, she only receives $1/H_{i,t}$ for each dollar of the frictionless price. Let us define a liquidity-adjusted return as

$$R^*_{i,t+1} = \frac{P_{i,t+1}/H_{i,t+1} + C_{i,t+1}}{P_{i,t}H_{i,t}}.$$

We can think of $H_{i,t}$ as one plus the fraction of the bond value that needs to be paid for the purchase of the bonds, due to a bid-ask spread. For simplicity, I assume that there is no liquidity concern for Treasury bonds. Then, applying the same log-linear approximation to $R^*_{i,t+1}$, I obtain the one-period identity for a log liquidity-adjusted excess return as

$$r^{e*}_{i,t+1} \equiv \log R^*_{i,t+1} - \log R^f_{i,t+1} \approx -\rho s\tau_{i,t+1} + s\tau_{i,t} - l_{i,t+1} - h_{i,t+1} + const,$$

where $h_{i,t+1}$ is a illiquidity measure defined by $h_{i,t+1} \equiv \rho \log H_{i,t+1} + \log H_{i,t}$. Iterating forward, I can obtain the three-way decomposition of the credit spread:

$$s\tau_{i,t} \approx \sum_{j=1}^{T-t} \rho^{j-1} r^{e*}_{i,t+j} + \sum_{j=1}^{T-t} \rho^{j-1} l_{i,t+j} + \sum_{j=1}^{T-t} \rho^{j-1} h_{i,t+j} + const. \tag{19}$$

This identity says that, holding discount rates and credit loss constant, higher illiquidity in the future leads to higher current credit spreads (lower current prices) for corporate bonds.

Taking the conditional expectation yields

$$s\tau_{i,t} \approx s\tau_{i,t}^{r} + s\tau_{i,t}^{l} + s\tau_{i,t}^{h},$$

where

$$s\tau_{i,t}^{h} \equiv E\left[\sum_{j=1}^{T-t} \rho^{j-1} h_{i,t+j} \middle| \mathcal{F}_t\right].$$

Therefore, we can decompose the variation of credit spreads into three components: changes in expected excess returns, expected credit loss and expected illiquidity. The question is how to measure $h_{i,t}$. There are variety of measures of illiquidity in the existing literature. For example, I could follow Bao, Pan and Wang (2011) in constructing the Roll (1984) measure to estimate the transaction cost. Specifically, for bond i, I compute

$$\sqrt{-Cov_t\left(\Delta p_{i,t,d-1}, \Delta p_{i,t,d}\right)},$$

where $\Delta p_{i,t,d}$ is the log price change between day $d-1$ and day d in month t. Few investors trade bonds every month. Indeed many of them hold the bonds until their maturity. To account for trading frequency, I can use $\sqrt{-Cov_t\left(\Delta p_{i,t,d-1}, \Delta p_{i,t,d}\right)}$ times the bond's turnover rate (monthly trading volume divided by the face value of the bond) to obtain the illiquidity measure $\log H_{i,t}$. Roll (1984) shows that the effective bid-ask spreads are $2\sqrt{-Cov_t\left(\Delta p_{i,t,d-1}, \Delta p_{i,t,d}\right)}$, when the fundamental value follows a random walk. Thus, $\log H_{i,t}$ measures the effective transaction costs for an investor whose portfolio has the average turnover rate.

Once I compute the illiquidity measure, $h_{i,t}$, I can estimate the VAR with the state vector $X_{i,t} = \begin{pmatrix} r_{i,t}^{e*} & s\tau_{i,t} & h_{i,t} & -\tau_{i,t}DD_{i,t} \end{pmatrix}'$ and infer the implied long-run forecasting coefficients. By comparing the volatility of $s\tau_{i,t}^{h}$ with the credit spread volatility, I can in principle quantify the contribution from the liquidity variation in explaining the credit spreads, controlling for risk premium and expected default.

Though this extension of the decomposition to include illiquidity is conceptually simple, it is not easy to empirically implement this three-way decomposition, due to the limited data availability. To my knowledge, the construction of any illiquidity measure requires the daily price and/or trading volume information. Since this information is available only in Mergent FISD/NAIC and TRACE, the sample period is limited to 1994 onwards. Even after 1994, the fraction of the bonds in these two databases is limited. This is problematic, as the model-free decomposition approach in this paper crucially depends on the large sample size. Since default occurs relatively infrequently, the time series dimension of the data must be sufficiently large so it covers at least several credit cycles. Also, the cross section must be large so it covers wider range of credit ratings. Thus, it is not feasible to reliably estimate the VAR with the subsample only from Mergent FISD/NAIC and TRACE.

Another issue in estimating the liquidity effect is a bias due to the sample selection. As I can only compute the liquidity measure for the bonds that are traded relatively frequently, the subsample based on Mergent FISD and TRACE is biased toward more liquid bonds. Thus, the resulting decomposition can only provide the lower bound for the role of liquidity.