Credit Spread Decomposition Notes

James Evans

06/17/2025

Abstract

The empirical evidence showing that a corporate bond's expected loss is only a small portion of a bond's credit spread is called the credit spread puzzle. This paper, using a reduced-form credit risk model, characterizes a risky bond's credit spread. This characterization provides a more general measure of a risky bond's credit risk and it shows that, in an arbitrage-free market, a bond's credit risk is only a fraction of the credit spread and not linearly related to the one-year, risk-neutral expected loss, resolving the credit spread puzzle.

The credit spread puzzle refers to the empirical observation that corporate bonds offer yields significantly higher than risk-free Treasury bonds, even when the expected loss from default is very small. To illustrate, consider a zero-coupon corporate bond that pays \$100 at maturity in one year. Suppose there is a 1% probability of default, and in the event of default, the bond pays nothing (i.e., zero recovery). Under the real-world measure $\mathbb{P}$, the expected payoff is:

$$\mathbb{E}^{\mathbb{P}}[\text{Payoff}] = 0.99 \cdot 100 + 0.01 \cdot 0 = 99$$

If investors were only compensated for expected default losses, this bond should trade at approximately \$99. In contrast, a risk-free Treasury bond maturing in one year would pay \$100 with certainty and—assuming zero interest rates—would also trade at \$100.

To address the credit spread puzzle, the paper adopts a reduced-form credit risk model in which defaults occur randomly with intensity $\lambda(t)$. This default intensity is not directly observable, but it can be estimated from market instruments like bond prices and CDS spreads. These prices embed information about perceived credit risk, allowing the model to extract implied default probabilities.

Using this framework, the authors define a new, more general measure of a risky bond's credit risk. Crucially, they show that even in an arbitrage-free market, a bond's credit risk accounts for only a fraction of its total credit spread. Moreover, the credit spread is not a linear function of the bond's one-year expected loss, which contradicts many prior modeling assumptions. This more general nonlinear relationship helps resolve the credit spread puzzle by explaining why credit spreads remain large even when expected losses are small.

Introduction

The determinants of risky bond credit spreads are an often studied empirical topic in finance, with still no consensus on the importance of the bond's expected loss in this decomposition. The low explanatory power of a bond's expected loss in explaining the credit spread is referred to as the "credit spread puzzle" (Feldhutter and Schaefer, 2018; Bai *et al.*, 2020). The approaches to analyzing the credit spread are one of the two: (i) based on a structural model for credit risk, to run a linear regression which decomposes the credit spread into various firm and market explanatory variables (e.g., Bai *et al.*, 2020; Campbell and Taksler, 2003; Collin-Dufresne *et al.*, 2001; Davies, 2008; Feldhutter and Schaefer, 2018; Huang and Huang, 2012); or (ii) based on a reduced-form credit risk model, to run a linear regression that characterizes the credit spread using the estimated default probabilities and other firm and market explanatory variables (e.g., Elton *et al.*, 2001; Giesecke *et al.*, 2011).

This section explains two major empirical approaches used by researchers to analyze the *credit spread puzzle*—the well-documented observation that corporate bond credit spreads are significantly larger than what would be implied by expected default losses alone. The primary goal of these studies is to understand what drives the size of the credit spread, defined as the additional yield that corporate bonds offer over risk-free Treasuries.

The **first approach** is based on a *structural credit risk model*, rooted in corporate finance theory, such as the Merton model. In these models, defaults arise when the market value of a firm's assets falls below a threshold defined by its liabilities. Researchers use firm-level data (e.g., leverage, volatility, asset value) to estimate default probabilities from this structural framework. They then run a *linear regression* of the form:

$$\text{Credit Spread}_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i$$

Here, X_{ij} represents explanatory variables derived from firm and market characteristics—such as leverage ratio, asset volatility, interest rate levels, or market-wide risk indicators like the VIX. The term ε_i represents the regression residual—the portion of the credit spread not explained by the included variables—capturing noise, omitted factors, or model imperfections. The regression attempts to quantify how much of the observed credit spread can be attributed to these factors. Examples of this approach include Bai *et al.* (2020), Campbell and Taksler (2003), and Collin-Dufresne *et al.* (2001).

The **second approach** uses a *reduced-form credit risk model*, in which default is modeled as a random event driven by a stochastic *intensity process* $\lambda(t)$, estimated directly from market prices—such as bond yields or credit default swap (CDS) spreads—without requiring information about the firm's balance sheet. From the estimated intensity, researchers derive default probabilities, then run a regression with the same goal: to explain the credit spread using the estimated probability of default and other control variables. The same linear regression framework applies:

$$\text{Credit Spread}_i = \beta_0 + \beta_1 \widehat{\text{DefaultProb}}_i + \beta_2 Z_{i1} + \cdots + \beta_k Z_{ik} + \varepsilon_i$$

where $\widehat{\text{DefaultProb}}_i$ is the market-implied probability of default and Z_{ij} includes additional firm-level and macroeconomic factors. Notable studies using this method include Elton *et al.* (2001), Giesecke *et al.* (2011), and Feldhutter and Schaefer (2018).

In both approaches, the core idea is to *decompose the credit spread* into weighted contributions from various explanatory variables using regression. This allows researchers to identify how much of the credit spread is attributable to actual credit risk (i.e., expected losses due to default) versus other sources—such as liquidity risk, risk premia, and macroeconomic uncertainty. The difference lies in how default risk is modeled: either structurally through firm fundamentals, or statistically through market-implied intensities.

> Unfortunately, structural model-based default probabilities have been shown in other contexts to be misspecified (e.g., Campbell *et al.*, 2008, 2011; Jarrow, 2011), calling into question the conclusions drawn from the structural approach to the credit spread decomposition. Also, the reduced-form credit risk approach is misspecified because it assumes a simple linear relation between the credit spread, default probabilities, and other market and firm explanatory variables. We show below that this relation is complex and nonlinear.

A model is said to be misspecified when its assumptions, structure, or included variables do not accurately reflect the true underlying process that generates the data. This can lead to biased estimates, incorrect inferences, or poor predictive performance.

The purpose of this paper is to provide a new decomposition of a bond's credit spread that resolves the credit spread puzzle and which can be used to understand the existing empirical evidence. This characterization, and the new credit risk measure derived herein, can be the basis for new empirical research on the determinants of a risky bond's credit spread.

To obtain this new decomposition, we use the risky bond valuation model contained in Hilscher *et al.* (2023). We show that the bond's credit spread is a nonlinear and complex function of the firm's default probability, recovery rate, risk premium, default risk premium, liquidity premium, and promised cash flows. This decomposition also provides a new credit risk measure for a risky bond, replacing the bond's one-year, risk-neutral expected loss. This new measure is needed because it is shown that the bond's credit spread is not a linear function of its one-year, risk-neutral expected loss.

This decomposition provides a theoretical explanation for the credit spread puzzle because: (i) the credit spread is not a linear function of the bond's one-year, risk-neutral expected loss; and (ii) the credit risk, although linear in this new credit risk measure, is still only a fraction of the credit spread. The remaining fraction corresponds to liquidity risk, which is a nonlinear function of the bond's default risk premium, liquidity premium, recovery rate, and the promised coupon and principal payments. It is important to emphasize that although this paper is theoretical, the risky debt pricing model from which the credit spread decomposition is obtained has been empirically validated with respect to the corporate bond market prices (Hilscher *et al.*, 2023). Consequently, this empirical validation implies that the credit spread decomposition provided herein is consistent with the observed and market determined credit spreads.

The *one-year risk-neutral expected loss* is the expected value of losses due to default occurring within the next year, as priced by the market. It is computed under the *risk-neutral probability measure* $\mathbb{Q}$, which reflects how investors price credit risk, including compensation for uncertainty and risk aversion.
Let the following variables be defined:

- F = face value (promised payoff) of the bond at maturity.
- R = recovery rate (fraction of F paid in the event of default), where $0 \leq R \leq 1$.
- τ = random default time.
- $T = 1$ = one-year time horizon.
- $\mathbf{1}_{\{\tau \leq 1\}}$ = indicator function equal to 1 if default occurs within one year, 0 otherwise.
- $\mathbb{E}^{\mathbb{Q}}[\cdot]$ = expectation under the risk-neutral probability measure $\mathbb{Q}$.

Then the **one-year risk-neutral expected loss**, denoted $\mathrm{RNEL}_{1\mathrm{yr}}$, is:

$$\mathrm{RNEL}_{1\mathrm{yr}} = \mathbb{E}^{\mathbb{Q}}\left[F \cdot (1-R) \cdot \mathbf{1}_{\{\tau \leq 1\}}\right]$$

Assume the following inputs:

- $F = 100$ (the bond promises to pay \$100 at maturity),
- $R = 0$ (zero recovery in case of default),
- $\mathbb{Q}(\tau \leq 1) = 0.04$ (the market-implied probability of default within one year is 4%).

Then the one-year risk-neutral expected loss is:

$$\mathrm{RNEL}_{1\mathrm{yr}} = 100 \cdot (1-0) \cdot 0.04 = 4$$

So, the expected loss under the risk-neutral measure is \$4. This means that if investors were only being compensated for default risk (and no other risks), the bond would trade at approximately:

$$\text{Price} = F - \mathrm{RNEL}_{1\mathrm{yr}} = 100 - 4 = 96$$

This paper resolves the credit spread puzzle by showing that the credit spread on a risky bond is not linearly related to its one-year, risk-neutral expected loss. Previous empirical models often approximated the credit spread using a scalar multiple of the expected loss under the risk-neutral measure. Formally, they assumed:

$$\text{Credit Spread} \approx \alpha \cdot \mathrm{RNEL}_{1\mathrm{yr}}$$

for some constant α. However, the paper demonstrates that this approximation is fundamentally flawed.

The authors define a more general measure of **credit risk** as the portion of a bond's price that is attributable to the possibility of default.

In this framework, credit risk enters the model *linearly*. That is, holding all else fixed, if the model-implied default intensity doubles, then the credit risk component of the bond price also doubles. This linearity refers to the way default-related inputs scale within the pricing equation of the model.

However, the paper emphasizes two key findings:

- The total credit spread is **not** a linear function of the one-year risk-neutral expected loss. The expected loss alone is a poor proxy for what actually drives spreads in the market.
- Even after defining credit risk precisely through their model, this credit risk component explains only a *fraction* of the observed credit spread.

The remaining portion of the credit spread is attributed to **liquidity risk**. This includes factors such as the bond's default risk premium, liquidity premium, recovery assumptions, and promised payments. Crucially, this liquidity risk enters the pricing relationship *nonlinearly* — meaning its effect on spreads is more complex and cannot be captured by a simple proportional change in input variables.