Correcting Biases in Corporate Bond Pricing Models: Risk Perception, Mispricing, and Arbitrage

Chang-Shu Chung*

Zicklin School of Business, Baruch College, City University of New York, USA

April 1, 2026

Abstract

Identifying mispricing exposures across systematic and idiosyncratic dimensions raises a nontrivial question: *how can these risks be effectively hedged out-of-sample?* This paper develops effective risk-management strategies for corporate bond portfolios and emphasizes the importance of correcting biases in structural pricing models caused by financial constraints. I propose two debiasing methods—incorporating risk perception and rank-difference pricing errors—to address structural issues after delta hedging, reducing portfolio risk by 46%, over 70% relative to standalone long positions. The information ratio rises from 1.23 based on raw spread errors to 2.95 achieving a 90% risk reduction. Despite limited arbitrage opportunities, effective stock–bond hedging depends on precise share management, with pricing-error signals guiding convergence and neutralizing structural risks to time temporary idiosyncratic mispricing.

JEL Codes: G11, G12, G13, G14, G32, E43, E44

Keywords: Capital structure arbitrage; Credit risk ranking; Bias correction; Nonlinearity; Structural risks; Leverage risk; Delta-hedging; Limits-to-arbitrage

---

*I am deeply grateful to my dissertation committee—Liuren Wu (Chair), Lin Peng, and Fabrice Tourre—for their guidance, insightful feedback, and unwavering support throughout my research. I also thank Linda Allen, Pierre Collin-Dufresne, Armen Hovakimian, Jian Hua, Margarita Tsoutsoura, Jun (Jonathan) Wang, Yajun Wang, Anastasia Zakolyukina, and Dexin Zhou for their valuable comments and discussions, as well as the seminar participants at Baruch College for their constructive feedback and engagement. All errors are my own. E-mail: changshu.chung@baruch.cuny.edu

1 Introduction

A central challenge for financial intermediaries in corporate bond investing is hedging systematic and idiosyncratic risks without diluting mispricing exposure along each dimension. Although structural models for pricing corporate bonds have been extensively developed to understand their valuation,[1] their role in guiding real-world investment decisions, particularly in *hedging*, remains an important and relatively less explored area. Building on the insights of Ross (1976) on the Arbitrage Pricing Theory (APT), which posits that asset prices are determined by their exposures to common systematic risk factors, this paper examines whether hedging these factor exposures enables investors to isolate idiosyncratic stock–bond mispricing. Specifically, this paper investigates *the extent to which systemic risk can be reduced through portfolio hedging strategies constructed using these models* and *whether such strategies effectively improve the information ratio by hedging the underlying risk*.

A longstanding challenge in this area is the *credit spread puzzle*, which arises from the inability of structural models to fully explain the average biases in actual credit spreads (Eom et al. (2004); Huang and Huang (2012)). Many studies have attempted to improve model fit through enhanced calibration techniques.[2] While these developments have advanced our understanding of various market risks and model performance, an important question remains regarding *how model-implied risks translate into portfolio-level investment outcomes.* In practice, investors manage portfolios consisting of multiple bonds and rely on structural models for both *valuation* and *hedging*. The growing complexity of bond markets, particularly since 2018, has driven a significant increase in bond portfolio trading volumes (e.g., Li et al. (2025)). This trend highlights the growing need for investors and

[1]For theoretical developments, see Merton (1974), Black and Cox (1976), Geske (1977), Leland (1994), Leland and Toft (1996), He and Milbradt (2014), Chen et al. (2018), and Feldhütter and Schaefer (2023), among others. Empirical evaluations of these models include Bharath and Shumway (2008), Schaefer and Strebulaev (2008), Bai and Wu (2016), Feldhütter and Schaefer (2018), Huang et al. (2025), and others.

[2]Previous studies have incorporated factors such as stochastic asset risk and asset volatility dynamics (e.g., Du et al. (2019)), alternative debt valuation approaches (e.g., Feldhütter and Schaefer (2018); Bai et al. (2020)), exogenous and endogenous liquidity (e.g., He and Xiong (2012); He and Milbradt (2014); Huang et al. (2025)), and time-varying macroeconomic conditions (e.g., Chen et al. (2018)).

market makers to accurately assess the risks inherent in diverse bond holdings (e.g., ICMA (2019)).[3] Effective portfolio construction, informed by a resolution of the credit spread puzzle, enables better mitigation of structural and underlying risks.[4][5] This paper explores the puzzle's role for bond portfolio management, offers new investment-oriented insights, and provides a clear, step-by-step framework for constructing bond portfolios that effectively hedge these risks.

The credit spread puzzle is fundamentally connected to *capital structure arbitrage*, in which investors seek to profit from valuation discrepancies between actual and model-implied spreads. Structural models that incorporate stock information offer a framework for assessing whether bond valuations deviate from equity valuations, a concept I refer to as "*firm-level relative valuation*." The term "*capital structure arbitrage*" is rooted in the theoretical framework of structural models, which models equity as a European call option and corporate bonds as covered options on the firm's assets, excluding equity. In theory, delta-hedging should eliminate most risk (e.g., Tian and Wu (2023)), suggesting that "*arbitrage*" implies the potential for positive returns with minimal risk exposure.[6] However, no exist-

[3]The International Capital Market Association (ICMA) notes that short selling allows market-makers to continuously quote prices for securities not held in inventory, supporting their role in financial intermediation. It also highlights the role of hedging: a long position in one security can be offset by a short position in a similar security, such that price fluctuations in one position are largely counterbalanced by opposite movements in the other. Through hedging, market-makers in secondary markets can manage interest rate risk arising from inventory holdings and temporary long positions acquired through trading activity (see ICMA (2019)). This paper aims to formally define "similar" assets as those exposed to common risks by developing a portfolio construction framework that aligns risk exposures across multiple dimensions, extending beyond interest rate risk to other relevant factors.

[4]Structural risks are a subset of systematic risks that affect all bonds but with heterogeneous sensitivities depending on bond characteristics, allowing the level of impact to be inferred from their structural features. In contrast, systematic risks that are not structural affect all bonds more uniformly or unpredictably.

[5]Jump-process models calibrated to historical data largely capture past behavior, but predicting future jumps remains inherently difficult. Since corporate bond markets are not fully efficient and mispricings are common, structural models offer a more practical framework for relative valuation — where explicit jump modeling adds limited incremental insight from a ranking perspective. Jumps stemming from both systematic and idiosyncratic sources can be managed through long-short positions that balance exposures in conjunction with the model's delta. This paper therefore prioritizes managing jump risk through hedging strategies rather than through explicit modeling alone.

[6]Firms issue or sell equity to obtain a premium from providing insurance to equity holders, and use the proceeds to repurchase debt as a hedge against underlying risks. Ma (2019) highlights a negative correlation between corporate financing flows across equity and debt markets, indicating that firms might act as arbi-

ing study fully achieves this arbitrage objective. This paper argues that the key limitations are *biases inherent in these structural models that distort valuation convergence* and *the presence of risks beyond those associated with the underlying asset*.

How can biases arise? These models assume that firms with similar inputs have similar levels of credit risk. This limitation can result in biased risk assessments. Small differences in creditworthiness are critical to distinguish between firms, but structural models often fail to capture these nuances. These models may underestimate risk for some firms while overestimating it for others, leading to inaccurate credit risk rankings for predicting price movements and, consequently, flawed hedging strategies.

I propose two simple debiasing methods—*cross-sectional* and *time-series debiasing*—on Merton (1974)'s model. Given the competitive nature of corporate bond markets in credit risk ranking, a more effective strategy may involve identifying overvalued or undervalued bonds through "*market-level relative valuation*," comparing bonds across firms. Motivated by the idea of gradual information revelation in Kyle (1985), this paper adopts a *cross-sectional debiasing* approach using *rank-difference*, assessing the relative valuation of actual versus model-implied credit spreads. Firms with initial ranking discrepancies tend to realign with their predicted positions in subsequent months, suggesting that relative mispricings are gradually corrected as market participants in both bond and stock markets incorporate available information.

In practice, this remains challenging because bond prices are exposed to structural risks, most notably *duration risk* and *credit risk*, as documented by Fama and French (1993). These risks influence numerous bond prices across the cross-section, though the magnitude of their impact differs. Although interest rate swaps can hedge the duration risk (e.g., Faulkender (2005)), their costs and complexities often hinder practical implementation. Similarly, managing credit risk by restricting investments to certain ratings is rarely fea-

---

trageurs. This paper examines hedging effectiveness for portfolio construction from a hedge fund manager perspective.

sible, as it reduces portfolio diversification.[7] When mispricing opportunities arise, hedge fund managers may prefer long-short bond portfolios that hedge both duration and credit risk exposures while capturing the associated relative-value opportunities. To effectively manage these risks across time, well-diversified portfolios are essential. This paper constructs quintile portfolios sorted by pricing errors. Firms with distinct characteristics tend to cluster in quintiles 1 and 5, enabling a long-short (quintile 5 minus quintile 1) portfolio that naturally offsets structural risks. The primary challenge lies in designing this diversification based on a well-defined and meaningful pricing error.

Ideally, pricing errors should be randomly distributed, exhibit diversity, and remain independent of firm-specific characteristics. To achieve this, I apply two debiasing techniques.[8] *Cross-sectional debiasing* involves ranking firms based on their credit spreads relative to peers, thus transforming raw values into ranks. I then compute the rank difference between actual and model-implied credit spreads, a process analogous to constrained linear regression with a zero intercept and a unit slope along a 45-degree line. The zero-intercept constraint eliminates systematic bias (i.e., debiasing), while the unit-slope constraint corrects for scale distortions, particularly the tendency of high-yield firms to exhibit disproportionately high spreads—a nonlinear pattern that the constraint helps control.[9]

However, the structural framework of the Merton model introduces a significant bias

[7] In the late 1990s, Long-Term Capital Management (LTCM) made aggressive investments in low-credit securities, such as junk bonds, based on the belief that market mispricings existed. Their strategy relied on high leverage and the expectation that credit spreads would narrow. However, during the global market volatility of 1998, LTCM's risky positions collapsed, resulting in one of the most notable hedge fund failures (see Lowenstein (2000)). This episode highlights the critical importance of managing structural risks.

[8] Almeida et al. (2023) propose a two-step methodology to improve the predictive accuracy of model-implied volatility surfaces. They first compute model-implied pricing errors, treated as biases, and then nonparametrically estimate the pricing error surface. The debiased model is constructed by adding this estimated error surface to the original model-implied volatility surface. This paper uses a similar framework but adds a discussion on why corporate bond pricing models are biased and addresses the issue of scale.

[9] Unlike linear or local fitting methods, the rank-difference approach accounts for the large scale difference in spreads between high-yield and investment-grade firms. Although model-implied spreads may be higher for high-yield firms, their actual spreads are disproportionately larger. As a result, portfolios based on linear or local fitting errors may mistakenly place high-yield firms in higher quintiles. The rank-difference method corrects for this by removing the influence of scale.

in assessing investment-grade firms.[10] When these firms issue substantial debt relative to equity, they may be incorrectly classified as high default risk, leading to firm-specific bias. To mitigate this, I implement *time-series debiasing* by conducting a firm-level linear regression of actual credit spreads, which reflect investors' risk perceptions, on model-implied spreads, using data from the prior half-year. This approach calibrates each firm's model output to a more appropriate relative ranking. The debiasing term is formed by adding the model's prediction to the historical average bias, which helps reduce the impact of firm-specific bias. In addition, this method provides a more accurate short-term forecast of credit risk movements by demonstrating the strong alignment between actual spreads and debiased model-implied spreads.

Subsequently, I apply the estimated coefficients to the model-implied spreads and find that their average rank correlation with actual spreads rises to 89% across all firms, compared to 66% without time-series debiasing. This strong cross-sectional fit suggests that models remain useful for relative valuation after debiasing, with residual mispricings likely driven by transitory, non-fundamental factors, potentially predicting future market movements (e.g., Bai and Wu (2016)). This two-dimensional debiasing approach mitigates the overrepresentation of high-yield firms in the higher quintiles and improves comparability between firms with varying creditworthiness, thereby enhancing portfolio diversification. By balancing pricing errors across investment-grade and high-yield firms, the portfolio composition remains stable. This ensures more consistent durations and credit ratings *across time* within each portfolio quintile.

Building on these insights, this paper outlines three important criteria for an effective capital structure arbitrage strategy: the first two address structural risks, and the third targets underlying risk.

The *first* criterion is that pricing errors cannot be predicted based on the bond's credit

[10]Here is another perspective on how structural models affect the credit risk estimation of investment-grade firms. The credit spread puzzle suggests that these models significantly underestimate spreads, particularly for short-maturity investment-grade bonds.

rating, maturity, and volatility, as bonds with lower credit ratings, longer maturities, and greater volatility typically exhibit higher yields (e.g., Wu and Zaman (2024)). This randomness creates greater diversity within quintile portfolios. The *second* criterion tests whether the standard deviations of the portfolio quintile returns exhibit consistency, which is influenced by the first criterion. The rationale is that a well-constructed long-short portfolio, formed by taking positions in the highest and lowest quintiles, should effectively hedge against these structural risks by equalizing the magnitude of their impacts. Thus, analyzing the distribution of firms within quintiles is crucial to determine whether the error structure aligns with these two objectives. The *third* criterion for evaluation is the effectiveness of delta-hedging, which measures how well the strategy reduces risk while preserving returns. The strategy takes a long position in the delta of unlevered bond returns and a short position in one-minus-delta of unlevered equity returns, effectively offsetting equity risk exposure. These three criteria are inherently interconnected, as effective delta-hedging must neutralize structural risks.

The use of unlevered returns (Choi and Richardson (2016); Doshi et al. (2019)) serves three critical purposes. First, it ensures methodological consistency with capital structure arbitrage implied by structural models, recognizing that bond valuation is contingent on the underlying asset rather than solely on equity. Second, it offers a *capital-structure-neutral* measure of risk exposure, reducing the influence of leverage on portfolio volatility. This enhances hedging effectiveness and enables more consistent comparisons across firms and portfolios. Third, it facilitates portfolio construction. I employ a mean-variance optimization approach, following Markowitz (1952), where portfolio weights are allocated based on the ratio of mean excess returns to covariance. While pricing errors provide predictive power for excess returns, their relationship with covariance is more complex. As discussed earlier, an optimal error structure improves diversification. To refine the allocation strategy, I assign greater weights to assets with lower covariance, which are typically associated with firms possessing lower default risk (higher delta) and relative high market values of

debt, thereby contributing to portfolio stability.

Empirical Results. All comparisons are out-of-sample: portfolios are formed at t and evaluated at $t+1$. The empirical results indicate that using the raw spread error to generate pricing error yields an information ratio of 1.23, with an excess monthly return of 0.32 and a portfolio standard deviation of 0.89. Applying cross-sectional debiasing using the rank difference error improves performance, generating an information ratio of 1.89, an excess monthly return of 0.27, and a portfolio standard deviation of 0.49. Implementing a delta-hedging strategy further reduces portfolio standard deviation to 0.45, increasing the information ratio to 2.06, corresponding to a 15.66% variance reduction. However, when limiting the analysis to investment-grade firms, delta-hedging reduces risk by only 5.19%, while for high-yield firms, the reduction is 48.29%. This suggests that, with only cross-sectional debiasing, delta-hedging is less effective for investment-grade than for high-yield firms.

Applying both debiasing techniques improves the information ratio to 2.52, with an excess monthly return of 0.28 and a portfolio standard deviation of 0.38. Delta-hedging further reduces risk by 45.71%, increasing the information ratio to 2.95 while lowering the portfolio standard deviation to 0.28. The high information ratios underscore the economic significance of forecasting bond and equity returns using relative valuations at the market level and firm level. In contrast, delta-hedging using only long positions reduces portfolio risk but also lowers the information ratio, resulting in less effective hedging strategies. This emphasizes that effective delta-hedging must neutralize structural risks. Moreover, the long-short strategy hedges approximately half of the portfolio's standard deviation, compared to 0.55 in quintile 1 and 0.62 in quintile 5, suggesting that structural risks beyond underlying risk are largely mitigated. Over 231 months (September 2002–December 2021), negative returns occurred in only 25 months post-hedging, with relatively small magnitudes, while cumulative returns increased steadily.

Examining delta-hedging performance, I find that it reduces risk by 40.49% for investment-

grade firms and 43.75% for high-yield firms. While cross-sectional debiasing helps mitigate risk, its impact is limited for investment-grade firms. In contrast, incorporating time-series debiasing provides additional benefits by further reducing portfolio risk. This improvement stems from calibrating firms to their correct relative rankings over time, which allows investors to more accurately identify mispricing opportunities. Hedge adjustments can be better timed, enhancing overall risk management. A key takeaway is that successful delta-hedging depends not only on precise share management but also on the ability to detect and respond to mispricing dynamics. I also implement two portfolio strategies based on the cross-sectional median of bond amounts outstanding, which serves as a proxy for liquidity (e.g., Longstaff et al. (2005)), consistent with prior literature documenting the role of liquidity in shaping credit spreads (e.g., Chen et al. (2007); Covitz and Downing (2007)). The portfolio with less liquid bonds achieves a higher information ratio of 3.64, compared to 1.38 for the more liquid portfolio. This suggests that greater liquidity enhances price efficiency, potentially reducing the profitability of value-based strategies. Evaluating performance over holding periods of two to eight months shows that delta hedging improves the information ratio for horizons up to seven months. Over these horizons, relative valuations gradually converge at both the market and firm levels, indicating that the identified idiosyncratic stock–bond mispricing is temporary.

The previous results are based on the literal implementation of the Merton model. To enhance bond valuation and hedge performance, I extend this framework by incorporating CDS-based credit spreads and hedging ratios, which help reduce redundancy in representing asset value and debt. The twisted approach reduces risk by 26.53% in all firms compared to literal setting, assuming a recovery rate of 0.7, increasing the information ratio to 3.01 and reducing the standard deviation from 0.28 to 0.24.[11] Compared to the unhedged

[11]The comparison examines portfolio risk through different valuation methods, acknowledging that variations in asset holdings could introduce biases when comparing them directly. However, it is important to recognize that these differences in asset choices reveal distinct risk profiles, which help assess the performance of the valuation methods in constructing errors. On the other hand, delta-hedging focuses on risk reduction within the same asset groups.

portfolio, the hedged portfolio shows a 60.11% drop in variance—outperforming the risk reduction in the literal setting.

For investment-grade firms, the risk reduction reaches 33.61% with a recovery rate of 0.8, improving the information ratio from 2.62 to 2.95 and decreasing the standard deviation from 0.27 to 0.22. For high-yield firms, the reduction in risk is 19.75% with a recovery rate of 0.7, increasing the information ratio from 2.24 to 2.50 and reducing the standard deviation from 0.48 to 0.43. These results highlight that time-series debiasing also further enhances the effectiveness of the twisted settings. Although the literal settings perform well in both valuation and hedging, the twisted settings achieve similar valuation accuracy with even better hedging performance. This suggests that effective delta-hedging benefits from less noisy model signals that reduce the noise inherent in linking the two markets, thereby improving the measurement of sensitivity between stock and bond prices.

Outline. The paper is organized as follows. Section 2 reviews the related literature and highlights the key contributions and gaps motivating this study. Section 3 presents the Merton model in both literal and CDS-based settings, introduces two debiasing approaches, and specifies three criteria for determining the effectiveness of capital structure arbitrage. Section 4 describes the data used as inputs to the structural models and the emergence of bias in the implementation of these models. Section 5 investigates the impact of abnormal yields on portfolio construction, evaluates performance through the information ratio and delta-hedging risk reduction, compares results under cross-sectional debiasing alone and in combination with time-series debiasing technique for both literal and CDS-based Merton models, and performs robustness checks using the Leland-Toft models. Section 6 concludes.

2 Related literature

This paper contributes to the literature on structural models by examining how financial constraints (e.g., Fazzari et al. (1988); Kaplan and Zingales (1997)) affect their estimation

of credit risk ranking (e.g., Bharath and Shumway (2008)). Firms with low credit ratings struggle to access external capital due to lenders' concerns about repayment ability, while firms with high credit ratings can easily obtain financing, enabling them to issue more debt even when their equity value is low (e.g., Rajan and Zingales (1998); Fama and French (2005); Sufi (2009)). This disparity introduces significant biases when estimating credit spreads for high-rated firms using structural models (e.g., Merton (1974); Black and Cox (1976); Leland (1994); Leland and Toft (1996)), as the leverage component in these models fails to accurately reflect the differences in firms' actual creditworthiness. This study contributes to the literature by demonstrating that, despite inherent biases and minor differences, structural models exhibit strong performance in valuation and hedging after the application of two debiasing techniques outlined in Section 5.4.

This paper also contributes to the literature on capital structure arbitrage strategies (e.g., Yu (2006); Duarte et al. (2007); Kapadia and Pu (2012); Chen et al. (2023)). Prior studies typically infer mispricing by comparing actual and model-implied credit spreads, but the frequent finding that actual spreads exceed model values does not justify a long-bond, short-equity position, as the trading signal's reliability remains unclear. Chen et al. (2023) highlights opposing return forecasts for stocks and bonds based on the debt-equity spread, though their analysis focuses on separate markets and leaves open questions about timing and hedging effectiveness. This paper advances the discussion by evaluating the practical hedging performance of the models, specifically its effectiveness in reducing risk, improving the information ratio, and guiding timely hedge adjustments.

Previous work evaluated hedge accuracy by comparing empirical debt-equity sensitivities with the hedge ratio implied by Merton (1974), beginning with Schaefer and Strebulaev (2008) and followed, for example, by Bao and Hou (2017), Choi and Kim (2018), Huang et al. (2020), and Huang and Shi (2021). Using two debiasing techniques, delta-hedging reduces portfolio risk by 45.71% in a long-short bond strategy, increasing the information ratio from 2.52 to 2.95 under monthly rebalancing based on one-month-ahead

unlevered return forecasts. These findings show that although the hedge ratio proposed by Schaefer and Strebulaev (2008) is highly effective, it must be applied with careful attention to biases in structural models, leverage risk, and associated structural risks beyond the underlying risk. To the best of my knowledge, this paper provides the first systematic demonstration of how hedge ratios derived from these structural models can be utilized in a stock–bond arbitrage framework to enhance the information ratio.

This paper also examines the corporate bond market through the lens of two existing research areas. First, it relates to research examining the effects of interest rates on credit spreads (e.g., Duffee (1998); Collin-Dufresn et al. (2001); Faulkender (2005); Gilchrist and Zakrajšek (2012); Choi and Kronlund (2018); Kuong et al. (2024); Li (2024); van Binsbergen et al. (2025)). Second, it relates to studies on how credit risk influences the credit spreads (e.g., Huang and Huang (2012); Friewald et al. (2012)). This paper shifts the focus to how pricing error measurements contribute to constructing portfolios with stable duration and credit characteristics across time. Wu and Zaman (2024) adjusts errors using local fitting, which partially reduces the standard deviation of the portfolio. Their approach addresses structural risks by examining average values in quantiles 1 and 5. This paper adopts a fundamentally different approach by introducing rank-based randomness, which enhances consistency with structural risks over time and significantly reduces variance, cutting the standard deviation of long-short positions by approximately 50% compared to portfolios in quantiles 1 and 5. This suggests that substantial structural risks are effectively hedged.

Building on these insights, this paper proposes that even a simple model, such as Merton (1974), can achieve strong performance in capital structure arbitrage by employing two debiasing techniques, potentially offering a potential resolution to the credit spread puzzle from an investment perspective. However, my results highlight several puzzles for practical application in investment. When constructing portfolio quintiles, the findings suggest that convergence signals are stronger for the highest quintile, which exhibits higher pos-

itive returns, while signals for the lowest quintile are weaker, with small absolute values that are either positive or negative. We would expect a higher negative return for a long position in the lowest quintile. This presents a puzzle. It not only dampens the performance of capital structure arbitrage but also raises questions about whether the valuation is inadequate for structural models when applied to firms with negative pricing errors, or whether short-selling constraints (e.g., Miller (1977); Duffie et al. (2002); Hong and Stein (2003); Vayanos and Weill (2008); Gromb and Vayanos (2010); Asquith et al. (2013); Bai and Collin-Dufresne (2019); Hendershott et al. (2020); Atmaz et al. (2024)), which limit the ability to correct overpricing.

Another key puzzle relates to market integration (e.g., Chen and Knez (1995); Duarte et al. (2007); Schaefer and Strebulaev (2008); Lin et al. (2011); Kapadia and Pu (2012); Friewald et al. (2014); Chordia et al. (2017); Choi and Kim (2018); Chen et al. (2018); Ma (2019); Sandulescu (2022); Kelly et al. (2023); Chen et al. (2023); Collin-Dufresne et al. (2024)). The results in this paper indicate that delta-hedging enables significant risk reduction and a high information ratio in a long-short bond portfolio by exploiting a robust measure of mispricing. This suggests that convergence signals in market-level relative valuation (Bharath and Shumway (2008); Bai and Wu (2016); Wu and Zaman (2024)) emerge quickly, while firm-level relative valuation between equity and bond markets (Ma (2019); Chen et al. (2023)) also adjusts rapidly, albeit only temporarily over a short horizon. Consequently, arbitrage opportunities remain limited due to imperfect and asymmetric signals of market integration (Sandulescu (2022)). Despite these constraints, the arbitrage strategy remains highly effective.

It is important to distinguish these limited arbitrage opportunities from the concept of limits-to-arbitrage. The latter refers to frictions, such as slow-moving capital (Duffie (2010)), that slow the market reaction to arbitrage opportunities, as discussed in the literature (e.g., Shleifer and Vishny (1997); Xiong (2001); Abreu and Brunnermeier (2002); Brunnermeier and Pedersen (2009); Gromb and Vayanos (2010); Kapadia and Pu (2012);

Cao and Han (2013); He and Krishnamurthy (2013); Brunnermeier and Sannikov (2014)). There is a puzzle in understanding whether the limited arbitrage opportunities are primarily from inefficiencies in information transmission between the stock and bond markets or from model inaccuracies, given that hedging with the underlying stock can effectively reduce the risk of the bond portfolio. Pursuing better resolutions to these puzzles is important, as it may lead to improved performance and more effective hedging strategies.

3 Model implementation

3.1 Merton (1974)'s valuation and hedge ratio

In Merton (1974), it is assumed that the total asset value (A) of a firm evolves according to a geometric Brownian motion, characterized by an instantaneous asset return volatility (σ_A). Furthermore, it is assumed that the firm has zero-coupon debt, with a promised payment of (D) at maturity (T). The equity (E) of the firm is conceptualized as a European call option on the firm's asset value, maturing simultaneously with the debt, and with a strike price equivalent to the face value of the firm's debt. The firm defaults if its asset value falls below the face value of the debt at maturity.

I can establish the following pair of equations that connect the firm's equity value and equity return volatility to its asset value and asset return volatility at time t as follows:

$$E = AN(d_1) - De^{-r\tau}N(d_2) \tag{1}$$

$$\sigma_E = (A/E)\sigma_A N(d_1) \tag{2}$$

where

$$\tau = T - t, \quad L = De^{-r\tau}/A, \quad d_1 = -\ln(L)/(\sigma_A\sqrt{\tau}) + 0.5\sigma_A\sqrt{\tau}, \quad d_2 = d_1 - \sigma_A\sqrt{\tau},$$

and r denotes the risk-free rate. I use the 1-month Treasury rate for r, as the focus is on cross-sectional differences between firms.[12]

This paper conducts the analysis in yield space, while preserving the term structure

[12] Alternative choices for r have negligible effects on cross-sectional performance. Bai and Wu (2016) use a zero interest rate for cross-sectional analysis.

effect in the cross-section. There are two main reasons for this choice. First, maintaining the variation in maturities allows for meaningful comparisons across firms, where longer-dated bonds generally carry higher risk due to greater exposure to interest rate and credit risk. Second, rather than matching each bond with a maturity-matched Treasury to compute spreads, I control for interest rate risk by constructing duration-matched portfolios. To maintain consistency with the terminology commonly used in credit risk models, I continue to use the term "credit spread" throughout the paper.

Equation (1) is the European call option valuation formula and links the firm's market capitalization to the value of the option. For Equation (2), I follow the method of Jones et al. (1984) and Bakshi et al. (2022) which uses the Itô formula to link the asset volatility to equity volatility. Starting with an initial guess on asset value and asset return volatility, I set the initial value of asset value as the sum of market capitalization and total book value of debt. The initial asset volatility is set as equity return volatility multiplied by the ratio of market capitalization to the initial asset value, i.e., $\sigma_A = \sigma_E(E/A)$. Then, I transfer Equation (1) and (2) to the following equations: (1) $D_1[A, \sigma_A] = -E + A[N(d_1) - LN(d_2)]$ and (2) $D_2[A, \sigma_A] = -\sigma_E + (A/E)\sigma_A N(d_1)$. The next step is to recover A and σ_A by minimizing the objective function: $f = \min_{A,\sigma_A} \left\{ [D_1(A, \sigma_A)]^2 + [D_2(A, \sigma_A)]^2 \right\}$. The primary objective of this estimation process, using these inputs, is to understand how the equity market conveys credit information and to compare relative valuations between the equity and bond markets. Next, I use the parameters to calculate the default probability, credit spread, and hedge ratio for each firm on a monthly basis.

The credit spread implied by the Merton model can be expressed as:

$$cs^{\text{MCSB}}(A, \sigma_A, D, t, T, r) = -\ln\big([N(d_2) + (1/L)N(-d_1)]\big)/\tau \tag{3}$$

$$cs^{\text{MCSC}}(A, \sigma_A, D, t, T, r, R) = -\ln\left(1 - \mathbb{Q}^{\text{Merton}}(A, \sigma_A, D, t, T, r)\right)(1 - R)/\tau, \tag{4}$$

Equation (3) provides a literal formulation of **M**erton's **C**redit **S**pread, derived from its market **B**ond value (B), termed **MCSB**. Equation (4) offers an alternative approach by computing the survival probability (i.e., one minus the Merton's default probability,

$\mathbb{Q}^{\text{Merton}}(A, \sigma_A, D, t, T, r) = N(-d_1)$) by assuming a constant hazard rate within the fair **C**DS spread framework, incorporating a recovery rate R, termed **MCSC**.[13]

For Merton's hedge ratio, I directly use $\partial B/\partial A$ because the underlying asset of the bond—modeled as a covered option (asset minus equity)—is the total asset value rather than just equity. Thus, unlike Schaefer and Strebulaev (2008), I adopt the following hedge ratios with respect to asset value:

$$h_B^{\text{MHRB}}(A, \sigma_A, B, D, t, T, r) = 1 - N(d_1), \tag{5}$$

$$h_B^{\text{MHRC}}(A, \sigma_A, B, D, t, T, r, R) = B(1-R)\phi(-d_1) \Big/ \left((1 - N(-d_1))\sigma_A A\sqrt{\tau}\right), \tag{6}$$

where ϕ is the probability density function of the standard normal distribution. Equation (5) represents the literal formulation of **M**erton's **H**edge **R**atio with respect to asset value, derived by differentiating the market value of the **B**ond, termed **MHRB**. Equation (6) provides an alternative hedge ratio by decomposing $\partial B/\partial A$ as $(\partial B/\partial CS)(\partial CS/\partial A) = (\partial B/\partial CS_1)(\partial CS_2/\partial A)$, where CS_1 is the literal credit spread given in Equation (3) and CS_2 is the credit spread from the **C**DS setting given in Equation (4), termed **MHRC**.

Intuitively, in the literal Merton model, information about asset value and debt is reflected twice in the market values of bond and equity, once through the payoff and once through the default probability. This setup helps reduce redundancy in representing asset value and debt. Several studies (e.g., Blanco et al. (2005); Longstaff et al. (2005); Ericsson et al. (2009); Bai and Wu (2016); Bai and Collin-Dufresne (2019); Bai et al. (2020)) use CDS spreads, which tend to yield relatively clean measures of credit risk. This paper focuses on corporate bond yields and highlights how the CDS-based model can offer more precise signals for valuation and hedging, providing new insights distinct from the existing literature.

[13] Bai and Wu (2016), Feldhütter and Schaefer (2018), and Bai et al. (2020) use this CDS framework.

3.2 Debiasing in time-series and cross-sectional analysis

In this paper, I use two sources of information to assess relative credit risk rankings: companies' accounting data within structural models and actual yields in corporate bond market. Structural models estimate firm-specific default probabilities and model-implied credit spreads based on accounting variables such as leverage, maturity, and asset volatility. These measures offer a theoretical perspective on credit risk rankings. Meanwhile, in the corporate bond market, investors trade bonds based on their risk perceptions, which are reflected in actual yields. By comparing these two rankings, I evaluate how well structural models align with the risk perceptions of market participants.

3.2.1 Debiasing in cross-sectional credit spreads

Directly subtracting the model-implied spread from the actual spread does not fully capture the mispricing opportunities discussed in Section 5.1, highlighting the need for a refined error adjustment. However, identifying a more effective error measure remains a challenge. Given that actual spreads often exceed model-implied spreads, particularly for high-yield firms, it is essential to refine the model to mitigate bias and account for scale differences.

Ideally, errors should be random, diverse, and independent of firm characteristics. To achieve this, I transform the raw values into ranks at time t, evaluating companies' credit risk based on their relative positions cross-sectionally. Higher credit spreads (both actual and model-implied) correspond to higher numerical ranks, with rank 1 indicating the lowest credit spreads. The rank error is defined as the rank difference:

$$d = \text{rank}(\ln(\text{Actual Credit Spread})) - \text{rank}(\ln(\text{Model-Implied Credit Spread})), \quad (7)$$

where I examine $\ln(\text{Model-Implied Credit Spread})$ with and without time-series debiasing in Sections 5.2.1 and 5.2.2. This is a cross-sectional comparison, and since both the actual and model spreads are adjusted by the same Treasury rate, the comparison takes place in yield space. For both actual and model-implied credit spreads, I apply a shifted approach by subtracting the minimum values to ensure they are non-negative, allowing the use of

the logarithm. Since the analysis focuses on ranking, this helps retain the data and does not affect the results.

Rank difference resembles linear fitting but with constraints: a zero intercept and a unit slope along a 45-degree line, potentially correcting for bias and scale, addressing the nonlinear relationship between actual and model spreads. This method identifies mispricing opportunities based on rank differences. I standardize the rank error by demeaning and dividing by its standard deviation:

$$\varepsilon_{\text{RankError}} = (d - \text{mean}(d))/\text{std}(d). \tag{8}$$

A positive rank error in higher quintiles suggests underpricing, indicating a long position, while a negative rank error in lower quintiles signals overpricing, suggesting a need to sell. Using these signals, I construct bond portfolios using $\varepsilon_{\text{RankError}}$ at time t and assess their monthly returns from t to $t+1$.

The key implications of cross-sectional debiasing using rank difference are as follows. Without rank difference, high-yield firms tend to concentrate in the highest quintile because their actual credit spreads are relative high. Although their model-implied spreads are also high, the magnitude of pricing error is larger than for investment-grade firms. Rank difference reduces the impact of scale, mitigating this issue. Additionally, since the intercept is zero, no systematic bias remains. This added randomness in quintile portfolios enhances their ability to hedge structural risks through long-short positioning.

3.2.2 Debiasing in time-series and cross-sectional credit spreads

For each firm, I perform a simple time-series linear regression over a six-month window[14] to correct for time-series bias in the model, using the historical actual credit spread. The time-series debiasing model is specified as:

$$\begin{aligned}&\ln(\text{Actual Credit Spread})_{t-1 \text{ to } t-s-1,i} \\ &= a_{t-1,i} + b_{t-1,i} \cdot \ln(\text{Model-Implied Credit Spread})_{t-1 \text{ to } t-s-1,i} + e_{t-1 \text{ to } t-s-1,i},\end{aligned}$$

[14]I also experimented with other time horizons, including one quarter, one year, and two years, but the results did not improve.

where s denotes the time interval and i is the firm index. Using the estimated parameters $\hat{a}$ and $\hat{b}$ at time $t-1$, I compute the model's time-series debiased prediction for t as

$$\hat{a}_{t-1,i} + \hat{b}_{t-1,i} \cdot \ln(\text{Model-Implied Credit Spread})_{t,i}. \tag{9}$$

This prediction is then compared to the actual credit spread at t using a debiased cross-sectional approach to determine the rank error, as discussed in Section 3.2.1.

The following are key implications for firm-specific time-series debiasing. Firms vary in their debt capacity and equity issuance. Investment-grade firms tend to have lower credit risk, but they can issue significant amounts of debt to support operations, which may lead to an overestimation of their credit risk when using a structural model based on accounting information. In contrast, high-yield firms have less flexibility in adjusting their debt and equity, making structural models more accurate in assessing their credit risk rankings. To address these discrepancies, I incorporate market perspectives, specifically, actual bond yields, to refine credit risk assessments, allowing market signals to inform credit risk rankings.

Time-series debiasing adjusts firms to their appropriate credit ranking, improving comparability. For instance, a structural model might misclassify an investment-grade firm with high debt issuance as having elevated credit risk. However, time-series debiasing corrects this distortion using market signals, allowing a fairer comparison with similar firms. If historical credit spreads and model-implied rankings align, the model may not improve credit risk mispricing detection. But if discrepancies exist, they indicate that the market and model offer different insights, potentially enhancing mispricing identification.

Cross-sectional debiasing based on rank differences helps mitigate both bias and scale distortions. However, due to inherent limitations in structural models, time-series debiasing is essential for aligning each firm with its appropriate credit ranking. By first correcting for time-series bias, firms can be properly ranked, enabling more accurate cross-sectional adjustments. The rank error is defined as the difference in ranks:

$$d = \text{rank}(\ln(\text{Actual Credit Spread}))$$

$$- \text{rank}(\ln(\text{Time-Series Debiased Model-Implied Credit Spread})). \quad (10)$$

The purpose of Sections 3.2.1 and 3.2.2 is to address the limitations of the model. Next, the discussion returns to the model to explore how to implement capital structure arbitrage, as Merton (1974) assumes that equity is a call option and bonds are covered options. This implies that hedging against the underlying asset should eliminate most of risk. The question arises: how can this be effectively implemented?

3.3 Capital structure arbitrage performance evaluation

The idea of capital structure arbitrage is to go long (short) on corporate bonds and short (long) stocks. This approach is similar to delta-hedging an option where the underlying asset is a stock. In the absence of significant arbitrage limits, dynamically hedging a call option with the underlying stock should eliminate most of the risk. The concept follows a similar logic: since a firm's equity functions as a European call option, and its bonds represent a covered option on the firm's assets minus equity, the hedge must account for the firm's total asset value rather than just its equity. Therefore, I hedge both equity and bond positions relative to the firm's total asset value.

To implement a delta hedge on equity, I hold one unit of equity and short delta units of the asset $(1 \times E - \text{delta} \times A)$, where delta is $\partial E/\partial A$. The resulting delta-hedged equity return is $\partial E/A - \text{delta} \times \partial A/A$, where $\partial E/A$ represents the *unlevered equity return*. Similarly, the delta-hedged bond return is $\partial B/A - (1 - \text{delta}) \times \partial A/A$, where $\partial B/A$ represents the *unlevered bond return*.

The scaling is based on the total firm value, allowing us to assess the true economic exposure of equity without the distortion caused by leverage. Leverage significantly amplifies equity volatility, as shown by Choi and Richardson (2016). Furthermore, Doshi et al. (2019) show that unlevered returns do not exhibit the nonlinear patterns found in levered returns, and that the volatility puzzle identified by Ang et al. (2006) applies only to levered returns. This distinction matters because pricing errors are used to predict future

price convergence, making it essential to isolate the effects of leverage. If leverage effects are not isolated, these predictions can be misleading. In other words, understanding the pure exposure to asset value (unlevered return) helps avoid confusing leverage-driven fluctuations. Therefore, leverage not only influences portfolio volatility and hedging effectiveness, but also the comparison of investment strategies.

Under the framework of Modigliani and Miller (1958), the asset value is determined by the total value of financial assets, indicating that firms with identical fundamentals but differing capital structures can have the same asset value. To compute the asset return $\partial A/A$, I begin by calculating the firm's market value A as the sum of the market capitalization and the market value of debt.[15] The market value of debt is the total debt multiplied by the weighted average bond price per dollar of notional value, denoted $\overline{B}$. I use the market value of total debt instead of just bonds, based on the assumption that investors price all of a firm's liabilities according to its overall credit risk, which is reflected in the prices of its tradable bonds.

Next, I convert the levered equity return $\partial E/E$ to the unlevered equity return, $\partial E/A =$ levered equity return $\times$ E/A. I use the market values of debt and equity from the data to compute the unlevered return. While one could use values estimated from structural models, this approach may introduce estimation errors, as highlighted by Doshi et al. (2019). Similarly, the levered bond return $\partial B/B$ is converted to the unlevered bond return, $\partial B/A =$ levered bond return $\times$ $\overline{B}/A$. Therefore, the asset return is given by $\partial A/A = \partial E/A + \partial B/A$.

The hedged return for equity is calculated as $\partial E/A - \text{delta} \times \partial A/A$ which simplifies to $(1 - \text{delta}) \times \partial E/A - \text{delta} \times \partial B/A$. To evaluate the delta-hedging performance for equity, I compare $\partial E/A$ with $(1 - \text{delta}) \times \partial E/A - \text{delta} \times \partial B/A$. Similarly, for delta-hedging performance on bonds, I compare $\partial B/A$ with $\text{delta} \times \partial B/A - (1 - \text{delta}) \times \partial E/A$.[16] Here,

[15] Choi and Richardson (2016) estimate the return on a firm's assets using a similar approach, but with three components: equity, bond, and loan.

[16] This paper focuses on structural models that allow for a closed-form expression of the hedge ratio. While

the hedge ratio differs from that in Schaefer and Strebulaev (2008), adopting the leverage-based formulation $(\partial B/\partial E)(E/B) = (1/(\partial E/\partial A) - 1)(E/B)$. This paper does not adopt this leverage-based hedge ratio and instead directly uses unlevered returns with delta to assess hedging effectiveness.

In the implementation, I calculate the portfolio returns using the weights of the market value of equity and debt relative to the market value of the total asset at time t. Specifically, I use the ratios E/A and $\overline{B}/A$ to assign weights to the stock and bond positions, which allows for a value-weighted approach to the portfolio. The returns are then calculated for the one-month period from t to $t+1$, considering the stock and bond performance over that time. With these adjustments, the criteria for evaluating capital structure arbitrage performance are now defined, and the next step is to implement strategies that minimize risk by identifying mispricing and applying hedging techniques to achieve arbitrage.

The primary challenge is identifying mispriced bonds to inform buy or sell decisions, as well as determining the optimal timing for hedging against risks arising from fluctuations in asset values. In corporate bond pricing structural models, market capitalization is used to compute a firm-specific relative valuation, which helps determine whether a bond is underpriced or overpriced relative to the company's stock price. When making buy or sell decisions based on this valuation, a delta hedge with the stock is required, as the bond's valuation is assessed in relation to that of the corresponding equity.

I present an illustrative example. Let's consider the convergence between market and model-implied bond prices. Suppose the market price of a corporate bond is $B_{\text{market}} = 92$ while the model-implied bond price, derived from the stock, is $B_{\text{model}} = 95$. Due to market inefficiencies, the market price is expected to converge toward the model-implied value; in this example, it partially converges to $B_{\text{new}} = 94$.

Chen et al. (2023) shows that the extended CreditGrades model effectively captures the valuation gap between debt and equity, it lacks a closed-form hedge ratio, making it less suitable for share management in hedging strategies. In contrast, this paper employs the Merton and Leland-Toft models, which provide closed-form hedge ratios, enabling clearer implementation and more robust hedging performance. Moreover, while Chen et al. (2023) emphasizes the information content of spread errors for insiders, this paper concentrates on practical portfolio construction for bond managers in hedge funds.

A deviation where B_{model} is higher than B_{market} indicates that the stock is overvalued relative to the bond. Suppose the stock price adjusts downward by $\Delta E = -20$. The corresponding change in the bond price can be calculated using the Merton model's delta. Let the bond delta with respect to assets be $\partial B/\partial A = 0.1$ which implies an equity delta of 0.9. The bond price change due to underlying risk is $0.9 \times \Delta B = 0.1 \times (-20) \Rightarrow \Delta B = -20/9 \approx -2.22$. This illustrates that bond prices are sensitive to stock movements — though in reality, the sensitivity may not always be strong.

From a mispricing perspective, buying the bond at $B_{\text{market}} = 92$ would suggest a potential gain of $94 - 92 = 2$ as it moves toward the model-implied convergence price. However, due to the underlying risk — the sensitivity of the bond to changes in total firm assets — the bond actually loses about 2.22, resulting in a net unhedged outcome of $2 - 2.22 \approx -0.22$, a loss despite the apparent mispricing.

To mitigate this risk, one can hedge by buying the bond and simultaneously shorting the stock. Since the bond and stock are linked through the firm's total assets, shorting the stock offsets the impact of underlying asset fluctuations on the bond price. In this example, the stock falls by 20, producing a profit of 2.22 from the short position, which exactly offsets the bond's loss due to underlying risk. After accounting for the hedge, the total profit equals the expected mispricing gain of 2. By hedging the stock exposure, an investor can capture the mispricing safely, effectively isolating the profit opportunity from asset-driven risk. This is a concrete example of capital-structure arbitrage, where the combination of positions in stock and bond exploits pricing deviations while controlling for the underlying risk embedded in the firm's assets.

This approach differs with the issue surrounding the credit spread puzzle, where the objective is to identify factors that explain the gap between observed and model-implied spreads. The focus here is on identifying instances of over- or under-valuation and detecting convergence signals between stock and bond prices. The signal for this relative valuation may be weak, reducing the necessity of trading and hedging these bonds with stocks.

Thus, it is important to assess *whether a significant signal exists between them*. Given the competitive nature of corporate bond markets, particularly in credit ranking, a more effective approach may be to identify overvalued or undervalued bonds using market-level relative valuation, comparing them across firms cross-sectionally. Subsequent hedging with stocks can then be used to evaluate the robustness of this strategy.

Once mispricing and hedging are identified, the next step is portfolio construction, which involves analyzing how pricing errors influence expected excess returns and risk (covariance). I use a mean-variance optimization approach, as proposed by Markowitz (1952), where portfolio weights are determined by the ratio of mean excess returns to covariance. However, a key question arises: *how does the model's pricing error translate into forecasts of mean excess returns and covariance?*

While pricing errors can help predict excess returns, their relationship with covariance is less straightforward. To improve portfolio performance, one approach is to increase the diversity of assets within each quintile, reducing variance and enhancing overall returns. Assets with lower covariance should be given greater weight in the portfolio. This can be achieved by allowing the model's pricing errors to vary randomly within each quintile and adjusting weights based on unlevered returns. Specifically, I do not assign different weights based on the size of pricing errors; however, using unlevered returns still results in assets receiving different cross-sectional weights for a *capital-structure-neutral* measure.

Since firms vary in size and leverage, comparing raw, levered bond returns across firms can be misleading, as it may increase the portfolio's exposure to leverage risk.[17] By scaling ∂B by the firm's total assets, summarized market values of equity and debt, I normalize bond price movements with respect to firm size. This adjusted measure indicates how much firm-specific risk bondholders are exposed to per dollar of assets. This normalization is particularly valuable for cross-sectional comparisons and portfolio construction,

[17] Choi and Richardson (2016) indicates that financial leverage amplifies equity volatility and that capturing the nonlinear relationship between leverage and stock returns is challenging. Similarly, higher leverage increases bond volatility by heightening default risk, making bond prices more sensitive to fluctuations in asset values (e.g., Collin-Dufresn et al. (2001)).

as it ensures more consistent evaluation regardless of a firm's capital structure.

Additionally, the unlevered bond return is multiplied by delta, which can be interpreted as the survival probability. A higher delta indicates lower default risk, often associated with larger firms and lower volatility. For example, bonds with a higher market value relative to their assets and a higher delta should receive more weight, while those with a lower market value and/or lower delta should receive less. Firms with a higher bond market value and lower default risk, typically large firms with lower volatility, tend to have lower covariance with other firms, making them more attractive for investment. Therefore, delta-hedging unlevered returns not only facilitate effective hedging but also help allocate appropriate weights within portfolio quintiles, improving overall performance.

Given the need for precise valuation and hedging techniques, as well as a robust strategy for portfolio construction, an effective capital-structure arbitrage model should exhibit the following criteria:

(1) **Evaluation Criterion 1 [Unpredictability of Pricing Errors]**: The pricing errors are not aligned with the bond yield ranking, which is usually influenced by credit ratings, maturity, and volatility. Typically, bonds with lower credit ratings, longer maturities, and greater volatility tend to have higher yields. However, the pricing errors should behave differently: they vary randomly and are not linked to these factors. This unpredictability in pricing errors means that mispricing cannot be forecasted based on the bond's credit rating, maturity, or volatility. The randomness in mispricing introduces greater diversity into the portfolio, reducing concentration risk by allowing a broader range of assets to be included.

If pricing errors were aligned with bond yield rankings, risk could become concentrated in specific quintiles. For example, high-yield firms typically exhibit both higher actual credit spreads and higher model-implied credit spreads. However, their pricing errors are also larger. To mitigate scale effects and enhance randomness, adjusting for rank differences is crucial. Otherwise, reliance on pricing errors may disproportionately place

high-yield firms in the highest quintile, reducing diversification.

(2) **Evaluation Criterion 2 [Consistent Risk Exposure Across Portfolios]**: Due to (1), portfolios constructed from pricing errors should differ significantly from portfolios constructed from bond yields. Sorting bonds by bond yields results in portfolios sensitive to duration and rating that introduce specific risks. Section 5.1 shows that while high yields are associated with higher returns, they also exhibit greater volatility. In contrast, sorting bonds based on pricing errors from a well-specified model should produce portfolios with similar risk exposures across quintiles—such as comparable ratings and durations—since variations in these characteristics correspond to different levels of risk exposure. For example, high-yield firms face greater credit risk compared to investment-grade firms. Accordingly, one objective is to enhance long–short portfolios by mitigating structural risks, such as interest rate fluctuations, that also affect corporate bond prices and yields. Consistent risk exposure across quintiles suggests that each quintile comprises a more balanced mix of firms with diverse risk profiles.

Without consistent risk exposure, high-yield firms may become concentrated in quintile 5, while investment-grade firms predominantly occupy quintile 1. Interest rate movements generally exert a stronger influence on quintile 1 and a weaker effect on quintile 5, introducing substantial interest rate risk into the long–short portfolio. In times of economic downturn, interest rate fluctuations may have a more pronounced effect on high-yield firms. Thus, substantial interest rate risk persists in the portfolio, regardless of the circumstances. By constructing a long-short bond portfolio with consistent risk exposure, one of the goals is to hedge against interest rate risk, eliminating the need for separate hedging tools like interest rate swaps. The long-short position itself directly manages these structural risks and allows us to focus on the impact of asset value fluctuations.

(3) **Evaluation Criterion 3 [Hedging Effects of Stock Price Dependence]**: From (2), maintaining consistent risk exposure allows us to account for structural risks while focusing on firm-specific risks. In Merton's framework, asset value underlies both equity and

bond, representing a combination of equity and bond. If bond valuation is highly sensitive to stock prices, hedging with stocks significantly impacts the portfolio. Conversely, if this relationship is weak, stock-based hedging has minimal effect.

This paper examines mispricing to evaluate signals of market convergence. If actual credit spreads are much higher than model-implied spreads derived from equity market information, it suggests that bond prices are undervalued while stock prices are relatively high. This mispricing signals a trading opportunity, going long on bonds and short on stocks, which also provides insight into firm-level valuation convergence. While this strategy helps address firm-specific risks from stock fluctuations, additional risks must be considered to improve hedging effectiveness. Consequently, (1) and (2) are essential for mitigating the structural risks that affect firms.

These three criteria are vital for formulating a strong capital structure arbitrage strategy using long-short bond portfolios. Prior to presenting the results, I will first describe the inputs that are linked to the models.

4 Data

4.1 Sample construction

In this paper, I analyze data from the following sources: the Center for Research in Security Prices (CRSP) for daily and monthly stock returns, Compustat for quarterly fundamentals, the Wharton Research Data Services (WRDS) Bond Database for monthly bond returns, and the Mergent Fixed Income Securities Database (FISD) for corporate bond characteristics.

I use the WRDS Bond CRSP Link Table to merge data from the WRDS Bond Database with CRSP. Then, I merge the firm's fundamental financial information from Compustat to the CRSP dataset using the CRSP/Compustat Link Table and assume that the fundamental statement information becomes available two months after the end of the firm's financial quarter.

I filter the bond data following the criteria in Bessembinder et al. (2009) and Dickerson et al. (2023). Eligible stocks must be listed on the NYSE, AMEX, or NASDAQ exchanges, with share prices of at least $5. I include only U.S. common stocks. Eligible bonds must be listed or traded in the U.S. public market, excluding those issued through private placements, under the 144A rule, Yankee bonds, non-U.S. dollar bonds, asset-backed bonds, and convertible bonds. Bonds must fall into one of three categories: U.S. Corporate Debentures ("CDEB"), U.S. Corporate Medium Term Notes ("CMTN"), or U.S. Corporate MTN Zero ("CMTZ"), with either fixed or zero coupons. I exclude bonds with a time-to-maturity of less than 12 months or more than 30 years, prices outside the range of $1 to $1,000, and insufficient trading information to calculate next-month returns.

For investment analysis, excluding bonds without transaction prices, and therefore no observable return in the following month, raises concerns about forward-looking bias. In practice, investors often contact dealers directly or obtain indicative prices from vendors to estimate valuations. However, for simplicity, this paper excludes these bonds from the sample. The holding of illiquid bonds may pose an inventory risk, but there are strategies to manage it. Since many corporate bonds trade on the over-the-counter market, investors can directly contact dealers or institutional investors to find potential counterparties. These private negotiations often facilitate trades even when a bond is not actively quoted on the screen. Additionally, the structure of the U.S. corporate bond market has evolved significantly. The global financial crisis of 2008–2009 acted as a catalyst for the adoption of electronic trading platforms, as market participants sought greater transparency and efficiency. This shift accelerated during the COVID-19 pandemic in 2020, when remote work increased the demand for digitized trading workflows.

Since the Merton model lacks the features necessary to generate precise pricing implications for individual bonds and can only provide a rough estimate of a firm's credit risk, I primarily aggregate bond-level monthly return, yield, coupon, credit rating, maturity, and duration to the firm-level using weighted averages based on the total face value

of each bond outstanding (e.g., Ma (2019)). The Merton model is then used to calculate the firm-level default probability, credit spread, and hedging ratio.

For this firm-level data, the final sample consists of 149,538 firm-month observations for 1,374 issuing companies, spanning from July 2002 to December 2021. The number of firms per month in my sample ranges from 461 to 720, with an average of 615.38. Table 1 presents summary statistics based on time-series averages of the cross-sectional distribution, providing a comprehensive overview of the data's key characteristics. The table reports the cross-sectional mean (Mean), standard deviation (SD), percentiles (1, 25, 50, 75, and 90), skewness, and excess kurtosis, offering insights into the distribution's shape and variability.

[Insert Table 1 here]

4.2 Bias onset and inputs in structural models

For Equation (1) and (2), the involved variables are: asset value (A), asset return volatility (σ_A), equity value (E), equity return volatility (σ_E), the face value of the firm's debt (D), debt maturity (T), and interest rate (r). The variables of asset value and asset return volatility require estimation using equations (1) and (2). However, the remaining variables can be directly tied to available data:

- Equity value (E): I calculate equity value from monthly market capitalization using the CRSP database by multiplying the monthly stock price (`PRC`) by the number of outstanding shares (`SHROUT`).
- Equity return volatility (σ_E): I calculate one-year realized daily stock return volatility using data from CRSP.
- Total debt (D): I use the quarterly total liabilities of a firm (`LTQ`) from Compustat.
- Debt maturity (T): I use the time-to-maturity (`TMT`) for each bond from the WRDS Bond Database. This is aggregated to the firm level using bond outstanding weights.

- Interest rate (r): I employ the 1-month treasury rate (`DGS1MO`) sourced from the Federal Reserve Economic Data.

- Recovery rate (R): I fine-tuned this hyperparameter by testing various values from a predefined range to compare the CDS-based model's performance.

Structural models, such as the Merton and Leland-Toft models, use a firm's capital structure, maturity, asset return volatility, default barrier, and coupon rate to estimate default probabilities, credit spreads, and hedging ratios. These models cannot differentiate between a firm's credit rating when firms share similar input values. As shown in Figure 1, for the top 20% of the spreads (higher credit risk category), there is a clear discrepancy between the model and the actual results. The model categorizes a large number of investment-grade firms in this group. In contrast, for the bottom 20% of the spreads (lower credit risk category), no significant differences are observed. This highlights that models treat firms with varying actual creditworthiness as having similar credit risk, which can lead to significant biases, especially for investment-grade firms. Consequently, it is necessary to apply the approach described in Section 3.2.2 to recalibrate firms with biased rankings to more accurate positions.

[Insert Figure 1 here]

Figure 2 further illustrates the implications of time-series debiasing, using Centex Corporation—an example of a firm with BB-BBB range credit ratings. The figure shows trends for four specific months: February 2006, May 2007, April 2008, and May 2009. For each of these months, a time-series regression is conducted using data from the previous six months. The regression estimates the intercept and slope, which are then used to predict the credit spread for the following month. The predicted credit spread from model is compared to the actual spread to assess the model's accuracy.

In February 2006, the time-series trend moves from the top-right to the bottom-left. Without time-series debiasing, the model's predicted spread deviates significantly from

the actual spread. However, when time-series debiasing is applied, the predicted spread aligns more closely with the actual observed spread, capturing the underlying trend accurately. This shows that time-series debiasing improves the comparability of model spreads with actual spreads. The debiased model spreads for May 2007, April 2008 and May 2009 also show significantly improved accuracy relative to the actual spreads, highlighting the consistent effectiveness of the time-series debiasing approach. One can use lagged actual spreads to predict current actual spreads. However, such trends may be unstable and can distort the identification of market-level relative valuation and, consequently, pricing errors. The model is therefore useful for stabilizing these trends by incorporating fundamental accounting and equity information.

[Insert Figure 2 here]

Figure 3 illustrates the rank correlations between actual credit spreads and those predicted by the literal Merton model setting using Equation (3), as well as the time-series debiased model using Equation (9). The rank correlation between the actual spreads and the original model's predicted spreads, which averages 66.4%. The rank correlation between the actual spreads and the time-series debiased model's predicted spreads, which averages 89.34%. The high average rank correlation does not contradict the findings of Kapadia and Pu (2012), which report low correlation between credit and equity markets due to limits-to-arbitrage and market segmentation. These limits prevent arbitrageurs from fully eliminating pricing anomalies. In my results, the rank correlation is high but still below one, indicating that such limits persist. However, the time-series debiasing process improves the model by adjusting firms, particularly investment-grade firms, to more accurate positions for comparing their credit risk, resulting in a higher correlation with the actual spreads.

Notably, the corporate bond market ranking adjusts quickly when the rank drops to lower values. For instance, the rank correlation for the red line dropped below 60% during the 2008 financial crisis, but later rebounded to 70-80% and eventually approached 90%.

After the 2008 crisis and before the 2020 COVID pandemic, the rank correlation remained above 90%, even reaching 95%. This suggests that credit rankings are a robust characteristic, even in the face of major macroeconomic disruptions. The rank correlation cannot reach one, indicating that some firms are mispriced, which results in incorrect rankings if the model is accurate and the market is not. We expect these mispriced firms to gradually move toward their correct positions over time. A high correlation, such as 90%, suggests that only a few firms remain misranked, while a low correlation, such as 70%, indicates that more firms may still be misranked. This raises the questions: will all firms eventually converge to the positions predicted by the model? Do lower rank correlations indicate weaker convergence signals? I further explore these issues through empirical investment analysis in Section 5.

[Insert Figure 3 here]

5 Empirical analysis

5.1 The influence of abnormal yields in portfolio construction

In this section, I first summarize the behavior of abnormal bond yields (e.g., Bessembinder et al. (2009)). Corporate bond returns exhibit predictable patterns, primarily driven by default risk, duration risk, and liquidity—factors that are reflected in credit spreads. As shown in Table 2, the first column lists quintile bond portfolios with long positions based on the actual credit spreads. The data reveal that as the actual spreads increase from quintiles 1 to 5, the monthly returns of the portfolio also increase. This trend highlights a key characteristic of bond yields: higher yields—driven by greater credit risk and/or longer duration—tend to generate higher future returns but also lead to increased portfolio volatility. This suggests that investors demand higher risk premiums for bearing an increased risk of credit and duration, as shown in Wu and Zaman (2024). The need for compensation for these risks has a significant impact on portfolio construction.

The second column shows the model-implied credit spreads based on the Merton model. Portfolio returns across different quintiles differ when using model-implied spreads compared to actual spreads. The mean return of the long-short portfolio is 0.34 with model spreads, compared to 0.58 with actual spreads.[18] However, the standard deviation of the long-short portfolio is lower with model spreads (1.11) than with actual spreads (1.39). This suggests that the model can reduce certain aspects of portfolio volatility. However, the model spread also follows a similar pattern to the actual spread.

To reduce portfolio volatility related to actual spreads while maintaining performance, I use the standard approach of calculating the raw spread error ($\varepsilon_{\text{Spread}}$), defined as the difference between the actual spread and the model spread, as

$$\varepsilon_{\text{Spread}} = \text{Actual Credit Spread} - \text{Model Credit Spread}.$$

in the third column. This method also aims to identify mispricing opportunities. Specifically, a spread error in higher quintiles suggests underpricing, while a spread error in lower quintiles indicates overpricing, signaling that bonds should be bought or sold accordingly. As shown in column three, spread error does not significantly improve portfolio performance. It only slightly reduces volatility in higher quintiles. However, in lower quintiles, such as quintile 1, portfolio volatility actually increases, with a standard deviation of 0.74 compared to 0.43 with actual spreads.

The key issue here is related to the well-known credit risk puzzle, where actual credit spreads are significantly higher than those implied by models. As shown in the last column, the spread error is 0.47 in quintile 1. However, the model-implied spread is 2.16, exceeding the 0.33 observed in quintile 2.[19] Based on the behavior of abnormal yields, taking long positions in these bonds generates higher returns and volatility. This contradicts the expectation that higher returns should come from shorting these bonds.

As discussed in Section 3.3, the pricing error must satisfy the criteria for effective cap-

[18] All portfolio constructions are evaluated using unlevered returns.

[19] Table 3 of Chen et al. (2023) presents similar results, showing that actual and model spreads exhibit a U-shape across spread quintiles.

ital structure arbitrage. The "Mean Error" column consistently increases from quintiles 2 to 5, and similarly, the "Model Mean" column rises over the same range. This suggests that the pricing errors are not random but are instead influenced by credit risk characteristics. This is further confirmed by the increase in the "Actual Mean" values from quintiles 1 to 5. The standard deviations of the portfolio increase from quintiles 2 to 5, with the value reaching 1.25 in quintile 5, indicating inconsistent risk exposure between portfolios. Consequently, using this pricing error directly does not constitute an effective arbitrage strategy, as a long-short bond portfolio based on this error generates a standard deviation of 0.89, which is higher than the 0.74 observed in quintile 1.

In other words, the spread error suffers from both bias and scale issues. This makes strategies based on quintiles 1 to 5 potentially ineffective. To address these issues, improving the measurement of spread error by considering rank differences can help adjust for these biases and scale problems.

[Insert Table 2 here]

5.2 Bias correction and scale adjustment

Since directly subtracting the model spread from the actual spread fails to capture mispricing opportunities, especially in the lower quintiles, it is crucial to find an approach to identify a better error measure. Because actual spreads are typically higher than model spreads, the primary concern is that cross-sectional fitting is heavily influenced by bias and scale. Addressing these issues requires a careful adjustment strategy.

5.2.1 Empirical analysis of debiasing in cross-sectional credit spreads

The first approach is to convert the values into ranks, i.e., evaluating companies' credit risk based on their relative rankings. I interpret the rank error in Equation (8) as a pricing error, calculated using the rank difference defined in Equation (7).

After correcting for cross-sectional bias and scale issues, portfolio performance based on pricing errors improves when sorted from quintile 1 to 5. As shown in the "Unhedged"

$(\partial B/A)$ column of Panel A in Table 3, the standard deviation of the long-short (5-1) position decreases from 0.89 for spread error to 0.49 for rank error, while the information ratio rises from 1.23 to 1.89. Additionally, the model spread exhibits an asymmetric U-shape across quintiles 1 to 5, with values of 0.72, 1.09, 1.59, 0.33, and 0.11. However, the high average actual spread of 4.71 in quintile 5 creates a squeeze, amplifying risk. This misalignment between model and actual spreads across quintiles 1 to 5 suggests that rank error nearly meets the first criterion, unpredictability of pricing errors, for effective capital structure arbitrage, compared to raw spread error.

The portfolio standard deviations—0.63, 0.74, 0.93, 0.82, and 0.85 for quintiles 1 to 5—are more consistent than those of spread error $(0.74, 0.62, 0.73, 0.86, 1.25)$. Implementing a long-short strategy reduces standard deviation to 0.49, which is lower than that of quintiles 1 and 5. This represents a risk reduction of approximately 40% compared to quintile 1 and 67% compared to quintile 5. By implementing the long-short delta-hedging bond portfolio on all firms using Equation (5), the hedge ratio from the Merton model reduces the portfolio standard deviation from 0.49 to 0.45, resulting in a variance reduction of approximately 15.66% as shown in the "Hedged" $(\text{delta} \times \partial B/A - (1 - \text{delta}) \times \partial E/A))$ column of Panel A. Additionally, the information ratio increases from 1.89 to 2.06, suggesting the effectiveness of this hedge ratio. This indicates that rank error meets the third criterion for a good capital structure arbitrage.

To satisfy the second criterion, the portfolio should consist of diverse firms with varying risk characteristics for maintaining consistent risk exposure across quintiles. Panel A appears to achieve this more effectively than the raw spread error approach. In the "Hedged" column, the standard deviation is 0.50 in quintile 1 and 0.71 in quintile 5, with the long-short portfolio showing a standard deviation of 0.45. This suggests that cross-sectional debiasing can help mitigate partial exposure to structural risks.

The main reason is model bias: high-rated firms are misclassified into higher ranks due to elevated model-implied credit spreads, while low-rated firms are misclassified into

lower ranks. As a result, using rank differences places low-rated (high-rated) firms into higher (lower) quintiles, increasing the standard deviation in higher quintiles and resulting in inconsistent risk profiles across quintiles. Section 5.2.2 later introduces time-series debiasing, which helps correct model biases, improves firm ranking accuracy, and ensures more consistent portfolio risk exposures.

Now, let's focus on investment-grade and high-yield firms to examine whether there is a difference. As shown in the "Unhedged" column of Panel B, the standard deviation of the long-short (5-1) position decreases from 0.70 for spread error to 0.38 for rank error, while the information ratio increases from 1.67 to 1.94. Additionally, model-implied spreads decrease from quintiles 1 to 5, whereas actual spreads exhibit the opposite pattern, increasing from quintiles 1 to 5. This misalignment suggests that rank error nearly satisfies the first criterion for effective capital structure arbitrage. The portfolio standard deviations—0.60, 0.62, 0.72, 0.75, and 0.78 for quintiles 1 to 5—are more consistent than those of spread error (0.48, 0.55, 0.68, 0.78, 1.05). Implementing a long-short strategy further reduces standard deviation to 0.38, which is lower than that of quintiles 1 and 5, demonstrating its effectiveness in managing structural risks. This supports the conclusion that rank error also nearly meets the second criterion for successful capital structure arbitrage. While the hedge ratio is effective, it reduces portfolio risk by only about 5.19% for investment-grade firms.

As shown in the "Unhedged" column of Panel C, the standard deviation of the long-short (5-1) position decreases from 1.11 for spread error to 0.89 for rank error, while the information ratio increases from 0.92 to 1.78. Additionally, model-implied spreads follow a random pattern from quintiles 1 to 5, while actual spreads increase. This misalignment suggests that rank error nearly meets the first criterion for effective capital structure arbitrage. However, the portfolio standard deviations—0.87, 1.14, 1.52, 1.19, and 1.47 for quintiles 1 to 5—are not more consistent than those of spread error (1.44, 0.97, 0.95, 1.01, 1.83). Implementing a long-short strategy reduces the standard deviation to 0.89, which is higher than that of quintile 1 but lower than that of quintile 5. As previously discussed,

satisfying the second criterion requires diverse firm representation in each quintile. Since high-yield firms appear in all quintiles, they remain exposed to structural risks affecting the entire high-yield segment. This suggests that rank error for high-yield firms does not meet the second criterion for successful capital structure arbitrage. However, it reduces portfolio risk by approximately 48.29%, indicating that cross-sectional debiasing alone is more effective at hedging underlying risk for high-yield firms.

[Insert Table 3 here]

From the discussions on investment-grade and high-yield firms, constructing a strong capital structure arbitrage bond portfolio requires a diverse set of firms with varying credit characteristics in each quintile. Without this diversity, the strategy becomes less effective, as delta-hedging performance weakens for investment-grade firms, and high-yield firms are more easily influenced by structural risks.

Furthermore, Panel A in Figure 4 shows that while the average durations (measured in years) have similar overall means, they are not consistently aligned throughout the period. As a result, the long-short portfolio fails to fully neutralize the impact of interest rate risk. Panel B indicates that firms with higher credit risk are concentrated in the quintile 5 portfolio, while those with lower credit risk are concentrated in quintile 1. This explains why the quintile 5 portfolio has a higher standard deviation (0.85) compared to quintile 1 (0.63) when unhedged. Even after hedging, this discrepancy persists. Therefore, using only the rank-difference method to measure error is not enough to satisfy the second criterion. This limitation further reduces hedging effectiveness, as it leaves structural risks unaddressed.

[Insert Figure 4 here]

5.2.2 Empirical analysis of debiasing in time-series and cross-sectional credit spreads

In this section, I address firm-specific bias by applying the method discussed in Section 3.2.2. After applying time-series debiasing to each firm, I recalculate the rank differences

and use the resulting rank errors to identify mispricing opportunities, as shown in Equation (10).

In Panel A of Table 4, time-series debiasing reduces the risk in the long-short bond portfolio from 0.49 (with only cross-sectional debiasing) to 0.38, while the information ratio increases from 1.89 to 2.52. I then apply delta-hedging, which lowers the portfolio's standard deviation from 0.38 to 0.28, achieving a 45.71% reduction in variance. The information ratio improves further, rising from 2.52 to 2.95. Notably, in quintile 5, the average actual (model) spread is 3.90 (0.67), lower than the 4.19 (0.73) in quintile 4. Quintile 5 is not the one with the highest spread, yet its portfolio returns are higher: 0.33 for "Unhedged" and 0.25 for "Hedged," exceeding the corresponding returns of 0.27 and 0.18 in quintile 4. The portfolio standard deviations are lower as well—0.78 for "Unhedged" and 0.62 for "Hedged," compared to 0.86 and 0.67 in quintile 4, respectively. These results suggest that the approach can enhance returns while reducing risk.

Furthermore, the portfolio standard deviations in the "Unhedged" and "Hedged" columns exhibit greater consistency. In the "Hedged" column, the long-short delta-hedged bond portfolio reduces the standard deviation by approximately 74.08% in quintile 1 (from 0.55 to 0.28) and by approximately 79.60% in quintile 5 (from 0.62 to 0.28). This suggests that debiasing time-series credit spreads improves both the effectiveness of long-short strategies in managing structural risks and the precision of delta-hedging in controlling underlying risk. While delta-hedging alone still reduces risk, it delivers lower information ratios without long-short positioning. This underscores the importance of accounting for structural risks when applying delta-hedging in structural models.

Panels C and D of Figure 4 further support this finding by showing smaller gaps between duration and credit rating in quintiles 1 and 5, contributing to the portfolio's stable standard deviation, which improves further with delta-hedging. Applying both cross-sectional and time-series debiasing aligns the pricing error with the second criterion, strengthening hedging effectiveness.

During the period from March 2020 to early 2021, bond portfolios in quintile 5 showed an average duration about two years longer than those in quintile 1, suggesting persistent interest rate exposure in the long-short portfolio. Since such gaps are rare outside this period, the anomaly may serve as a reference point for investors assessing interest rate risk.

In Panels B and C of Table 4, the hedge ratio exhibits a more effective pattern compared to the results in Panels B and C of Table 3. It reduces long-short portfolio risk by approximately 40.49% for investment-grade firms and by about 43.75% for high-yield firms. This suggests that combined time-series and cross-sectional debiasing are highly effective at hedging underlying risk for both investment-grade and high-yield firms. This method reduces the risk from 0.89 (spread error) to 0.28, achieving a variance reduction of approximately 90% in these two valuation approaches.

In Panel B, the long-short delta-hedged bond portfolio for investment-grade firms reduces standard deviation by approximately 74.05% in quintile 1 (from 0.53 to 0.27) and 79.06% in quintile 5 (from 0.59 to 0.27). In Panel C, for high-yield firms, the standard deviation is reduced by approximately 51.60% in quintile 1 (from 0.69 to 0.48) and 71.56% in quintile 5 (from 0.90 to 0.48). These results suggest that the delta-hedged, long-short strategy—especially when combined with time-series debiasing—successfully removes a large portion of the structural risks in corporate bond returns. In simpler terms, this method allows us to focus on excess returns that are not driven by known structural risks.

As discussed in Section 3.3, the pricing error meets the criteria for effective capital structure arbitrage. In the first column of the "Pure Model" in Panel A, the average model-implied credit spreads increase from quintiles 1 to 3 and then decline in quintile 5. A similar pattern is observed in the average actual credit spreads. In contrast, the "Debiasing Model" shows no clear trend, suggesting randomness. These patterns support the first criterion for pricing error. In the "Unhedged" column, portfolio standard deviations increase from 0.65 in quintile 1 to 0.95 in quintile 3, then decline to 0.86 in quintile 4 and 0.78 in

quintile 5. This indicates not only more consistent risk exposure across portfolios for satisfying the second criterion but also a reduced level of risk when underpricing opportunities appear in quintile 5.

[Insert Table 4 here]

Although rank correlations dropped to between 0.5 and 0.8 during the 2008 financial crisis, as shown in Figure 3, Panel B of Figure 5 demonstrates that long-short bond portfolios still generated positive returns, suggesting that the signal for market bond yields converging to model predictions remained strong. In contrast, during the 2020 COVID-19 pandemic, the rank correlation fell below 70%, accompanied by negative portfolio returns, a pattern distinct from the 2008 financial crisis. While the rank correlation quickly rebounded to 90% with positive portfolio returns, the sharp negative returns during the pandemic raise an important question: *What external risks were reflected in the portfolio despite hedging?*

A high rank correlation does not necessarily indicate a strong convergence signal, highlighting the importance for investors to closely monitor market dynamics and government interventions. For example, while the rank correlation exceeded 90% in February 2020, the monthly unlevered return of the long-short bond portfolio dropped below -1% between February and March 2020. Although the portfolio returns rebounded quickly in the following months, this sharp decline illustrates that a high rank correlation alone cannot fully capture sudden market disruptions.

In response to the financial turmoil caused by the COVID-19 pandemic, the Federal Reserve announced the establishment of the Secondary Market Corporate Credit Facility (SMCCF) on March 23, 2020, to stabilize the corporate bond market and ensure liquidity (e.g., Li and Ringgenberg (2024)).[20] By June 2020, the SMCCF began actively purchasing a diversified portfolio of corporate bonds, helping to restore market confidence and improve credit availability for large investors. Despite these stabilizing efforts, the increased

[20]See: https://www.federalreserve.gov/monetarypolicy/smccf.htm

intervention also led to heightened portfolio volatility, reflecting the complex interplay between market dynamics and policy execution. As noted in Crouzet and Tourre (2021), such credit interventions can be counterproductive during downturns by exacerbating debt overhang and dampening long-term investment. In contrast, the results presented here suggest that the SMCCF may have promoted greater alignment in market-implied credit rankings and facilitated value convergence. Consequently, this policy may offer benefits to value-oriented investors.

In addition, effective capital structure arbitrage requires neutralizing structural risks and implementing delta-hedging, both of which involve substantial shorting of bonds and stocks. This underscores how short-selling restrictions (e.g., Bris et al. (2007); Beber and Pagano (2013); Grullon et al. (2015); Goyal et al. (2025)) may impair bond managers' ability to manage both underlying and structural risk exposures.

Overall, Panel B of Figure 5, which employs both cross-sectional and time-series debiasing techniques, indicates that, over the 231-month period from September 2002 to December 2021, negative returns occurred in only 25 months following hedging, and their magnitudes were relatively modest.[21] By contrast, Panel A, which employs only cross-sectional debiasing, exhibits more frequent negative returns and less effective hedging. Collectively, these results suggest the presence of a low-risk arbitrage opportunity once key risks are effectively controlled, although such opportunities are inherently limited, as delta-hedging does not invariably enhance portfolio returns and may, in some instances, merely yield breakeven outcomes.

Moreover, hedging with stocks is especially effective when the long-short bond portfolio experiences negative returns. In such periods, losses in the bond portfolio can be partially offset by gains in stocks, reducing overall portfolio losses. Even when bonds are undervalued, the total market value of assets may still decline, particularly during crises,

[21] These findings differ from those reported in Chen et al. (2023)'s Figure 2, which documents a pronounced decline in both stock and bond portfolios during the 2008 financial crisis. A possible explanation is that the combined debiasing approach mitigates both credit and duration risks, which the spread error alone cannot fully neutralize over extended periods.

often driven by falling stock prices. This highlights the importance of including stocks in the hedging strategy. The resulting counterbalancing effect helps stabilize returns and enhances the effectiveness of hedging during adverse bond market conditions.

[Insert Figure 5 here]

Liquidity in corporate bond markets has been shown to significantly impact credit spreads and, consequently, the speed and extent of convergence signals.[22] This effect does not notably reduce investment performance across all firms in the sample. An interesting question arises: *do bonds with higher or lower liquidity generate greater profits for investors?* On the one hand, higher liquidity allows for easier and faster trading, but it may limit value-investing opportunities because bond prices tend to adjust quickly to reflect their fair value. This rapid convergence can reduce the potential to earn excess returns through mispricing. On the other hand, lower liquidity might offer more pricing inefficiencies, and thus greater opportunities for value investing. However, it also introduces higher trading risks, such as wider bid-ask spreads or difficulty in executing large trades.

To examine this trade-off between liquidity and profitability, I assume that the bond manager's trading behavior is guided by the amount of bonds outstanding, which serves as a proxy for market liquidity (e.g., Longstaff et al. (2005)). Specifically, I assume that the bond manager uses the cross-sectional median of bonds outstanding to form two strategies: one based on firms with above-median bond issuance (higher liquidity) and the other on firms with below-median issuance (lower liquidity).

[Insert Table 5 here]

The results are shown in Table 5. Panel A reports results for bond portfolios constructed from firms with above-median bond amounts outstanding. The long-short strategy in the "Hedged" column yields an information ratio of only 1.38, significantly lower

[22] See, for example, Chen et al. (2007); Covitz and Downing (2007); Brunnermeier and Pedersen (2009); Lagos and Rocheteau (2009); Bao et al. (2011); Lin et al. (2011); Friewald et al. (2012); Dick-Nielsen et al. (2012); He and Xiong (2012); Acharya et al. (2013); He and Milbradt (2014); Bessembinder et al. (2018); Chen et al. (2018); Goldstein and Hotchkiss (2020); Goldberg and Nozawa (2021); Huang et al. (2023).

than the 3.64 observed in Panel B, which is based on firms with lower liquidity. This suggests that higher liquidity corresponds to a more efficient market, where prices quickly incorporate available information. As a result, value-investing strategies based on monthly rebalancing have less opportunity to earn excess returns because mispricings are corrected faster. In contrast, lower liquidity markets tend to be less efficient, with slower price adjustments and more persistent pricing anomalies, which can create greater profit opportunities for value investors. The rank correlations remain high in both cases—87.13% for high-liquidity firms and 89.13% for low-liquidity firms—indicating consistent signal rankings. However, the strategy is less profitable in high-liquidity environments. In short, liquidity improves price efficiency, which can reduce the profitability of value-based strategies; however, delta-hedging remains effective for both types of liquidity-based strategies, indicating that the two debiasing techniques continue to be robust tools for identifying mispricing, regardless of liquidity conditions. The findings also align with previous research suggesting that investors require a liquidity premium to hold illiquid securities (Amihud and Mendelson (1986)). In both unhedged and hedged strategies, bond portfolios with lower liquidity generate higher excess returns than high-liquidity portfolios in quantiles 3, 4, and 5.

Motivated by the limits-to-arbitrage literature, pricing discrepancies are expected to persist but gradually decline over longer horizons, reflecting slow price convergence caused by frictions and arbitrage constraints. By constructing a pricing error that satisfies three key criteria for capital structure arbitrage, I form long-short bond portfolios with relatively low risk. To assess the consistency of this strategy with the literature, I examine holding period returns across different horizons (Jegadeesh and Titman (1993)). Table 6 analyzes portfolio holding returns from two to eight months to evaluate whether arbitrage opportunities persist and if effective hedging remains viable over longer periods. Specifically, portfolios are constructed each month but held for k months, where k ranges from 2 to 8. Reported returns are the average monthly returns over the holding period. The re-

sults show that delta-hedging in long-short bond portfolios effectively reduces risk and improves the information ratio for holding periods of two to seven months. These findings suggest that arbitrage opportunities persist for up to seven months and that delta-hedging remains effective for longer holding periods once structural risks are addressed. The results also indicate that relative valuations continue to converge at both the market and firm levels over longer—though not excessively long—horizons, indicating that the identified idiosyncratic stock–bond mispricing is temporary.

[Insert Table 6 here]

To test the robustness of these findings, I develop a trading strategy where long-short bond portfolios are held for a maximum of six months, with positions maintained if monthly returns remain negative. In Figure 6, the hedged portfolios achieve an information ratio of 3.67, up from 3.10 without hedging, while the portfolio standard deviation falls from 0.34 to 0.23, a 54.24% reduction in variance. Across 231 trading months, only 9 exhibit negative returns, all of which are small in magnitude, suggesting that the strategy functions as a low-risk arbitrage with stable performance. These results highlight the persistence of valuation convergence or divergence between the bond and stock markets. The model maintains strong out-of-sample performance across two major market disruptions — the 2008 financial crisis and the 2020 COVID-19 pandemic — indicating that rank-difference error construction paired with time-series bias correction robustly manages structural risks.[23]

[Insert Figure 6 here]

[23]Transaction risk, arising from the inability to execute trades due to incomplete price information, is not addressed in this study.

5.3 Debiasing in time-series and cross-sectional analysis of twisted credit spreads and hedge ratios

In the previous section, I focus on using the literal version of the Merton model for credit spreads and hedge ratios. Although the results show strong performance for bond portfolios, these settings present an issue: asset value and debt are included twice in the formulas. The hedge ratio for a bond measures how changes in the asset value relate to changes in the bond value. Including both asset and debt information introduces complexity.

In this section, I relax this assumption and use a CDS-based framework to adjust the credit spread and the hedge ratio, as outlined in Section 3.1. I use Equation (4) for the credit spread, and Equation (6) for the hedge ratio. Compared to the literal setting, these formulas introduce an additional parameter, R, the recovery rate. I explore the impact of different recovery rates ($R = 0.5, 0.6, 0.7, 0.8$) in Table 7.[24]

[Insert Table 7 here]

Panel A of Table 7 shows that for all firms, the model with $R = 0.7$ performs best, achieving an information ratio of 3.01, compared to 2.95 in the literal setting. The portfolio standard deviation decreases from 0.28 to 0.24, representing a 26.53% reduction in variance. Comparing the hedged portfolio with the unhedged one, we observe a 60.11% reduction in variance, outperforming the 45.71% risk reduction observed in the literal setting. Since the valuation performance is similar to the literal setting, this variance reduction is primarily driven by the CDS-based hedge ratio. This suggests that the CDS-based hedge ratio provides more precise information for effective hedging. However, cross-sectional debiasing alone does not improve performance in this CDS-based framework. In contrast, combining cross-sectional and time-series debiasing further reduces risk, underscoring the

[24]The use of a fixed 40% recovery rate has become a standard simplification in the CDS literature (e.g., Bai and Wu (2016); Huang et al. (2020)). Some papers (e.g., Feldhütter and Schaefer (2018); Bai et al. (2020)) instead use 37.8%, reflecting Moody's average historical recovery rate. This paper takes an investment-oriented approach by selecting the recovery rate that achieves the best performance in valuation and hedging, rather than relying on historical averages.

critical role of time-series debiasing in improving hedging precision.

In Panel B, for investment-grade firms, $R = 0.8$ yields the best performance, increasing the information ratio from 2.62 in the literal setting to 2.95. The portfolio's standard deviation decreases from 0.27 to 0.22, corresponding to a 33.61% reduction in variance. The long-short delta-hedged bond portfolio further reduces standard deviation by approximately 81.39% in quintile 1 (from 0.51 to 0.22) and 84.00% in quintile 5 (from 0.55 to 0.22), suggesting that the CDS-based setting offers more consistent risk exposure across quantiles and more effective neutralization of structural risks.

Without time-series debiasing, CDS-based models may fail to reduce risk and can even increase it, leading to lower information ratios. This reinforces the importance of time series debiasing in improving valuation accuracy and hedging effectiveness, especially for investment-grade firms. One reason is the volatility of accounting-based inputs. By applying time-series debiasing with historical actual spreads, we can more accurately rank credit risk and improve hedging performance. This approach is particularly effective in CDS-based settings, where it offers an advantage over the literal model by reducing the reliance on noisy information.

Panel C examines high-yield firms and finds that the model with $R = 0.7$ performs best, increasing the information ratio from 2.24 (in the literal setting) to 2.50. The portfolio standard deviation decreases from 0.48 to 0.43, a 19.74% reduction. The long-short delta-hedged bond portfolio further cuts standard deviation by approximately 65.30% in quintile 1 (from 0.73 to 0.43) and 77.67% in quintile 5 (from 0.91 to 0.43). This suggests that CDS-based models offer slightly better protection against structural risks for high-yield firms compared to literal models.

Unlike investment-grade firms, where accounting information may misrepresent credit risk and rankings, high-yield firms benefit more from cross-sectional debiasing, as their accounting inputs better reflect actual rankings. As a result, hedging is more effective. Nevertheless, combining cross-sectional and time-series debiasing still further enhances

the model's ability to detect mispricing and improve hedging performance for high-yield firms.

Overall, the findings show that CDS-based models are more effective than literal models at controlling structural risks and lowering portfolio risk through delta-hedging. Models without time-series debiasing may fall short, especially for firms with biased accounting measures. Combining both debiasing techniques improves performance further. These findings highlight the importance of less noisy model signals for effective delta-hedging, as they reduce noise in the connection between the stock and bond markets and improve the measurement of price sensitivity. Moreover, investment-grade firms with lower credit risk are more likely to undergo restructuring rather than liquidation, leading the bond market to implicitly assign higher expected recovery rates (e.g., Antill and Grenadier (2019); Antill (2022)). Accordingly, models must incorporate this market-implied recovery structure to generate accurate valuation signals and support effective hedging strategies.

5.4 Robustness check using various structural models

In this paper, I focus on addressing the bias and scale issues in the Merton model when applied to capital structure arbitrage. Over the past five decades, structural models have evolved significantly. Prominent extensions, including Black and Cox (1976) and Leland and Toft (1996), build on Merton (1974)'s framework. These models aim to better capture corporate policies related to default triggers before maturity, and coupon payments, in order to obtain a more accurate fundamental valuation. An interesting question that naturally follows is: *To what extent do these additional features help mitigate, or possibly influence, the bias observed in the original Merton framework?*

This section examines a version of the Leland-Toft model, which extends Black-Cox model with a constant proportional default barrier[25] and by incorporating continuous coupon payments.[26] The goal is to assess whether these modifications improve relative val-

[25]See Feldhütter and Schaefer (2018), Bai et al. (2020), and Feldhütter and Schaefer (2023)

[26]I do not include settings for the tax rate and payout ratio because key extensions—a default barrier before

uation performance for cross-sectional ranking and strengthen convergence signals. Following Merton, I derive both the literal and CDS-based settings for Leland-Toft's credit spreads and hedge ratios as shown in Appendix 7.1.

Figure 7 compares the rank correlation between the Merton and Leland-Toft models. Panel A shows that the rank correlations in the literal Leland-Toft settings are lower than in the Merton model (66.32%) when using default boundaries at d of 0.1 (56.75%), 0.3 (56.79%), 0.5 (56.77%), and 0.7 (56.43%) on average. Panel B demonstrates the rank correlation between time-series debiased literal model-implied credit spreads and actual spreads. The rank correlations approach the Merton model (88.86%) when using default boundaries at d of 0.1 (88.26%), 0.3 (88.26%), 0.5 (88.21%), and 0.7 (88.21%) on average. This suggests that imposing a fixed default boundary and continuous coupon payments in the literal setting does not significantly reduce bias, yielding results similar to those of the Merton model. A fixed default barrier may be too strong an assumption, as firms have dynamic default boundaries influenced by various factors over time.

Panel C shows that the rank correlations in the CDS-based Leland-Toft settings with R of 0.7 are lower than in the Merton model when using default boundaries at d of 0.7 (64.99%), 0.8 (65.17%), 0.9 (65.09%), and 1.0 (52.84%) on average. Notably, the rank correlations for d of 0.9 can exceed those of the Merton model on certain dates, suggesting that reducing redundant information may help the Leland-Toft model better capture relative rankings. Panel D demonstrates the rank correlation between time-series debiased CDS-based credit spreads and actual spreads. The rank correlations remain very close to the Merton model when using default boundaries at d of 0.7 (88.73%), 0.8 (88.72%), 0.9 (88.82%), and 1.0 (88.91%) on average. This suggests that imposing a fixed default boundary and continuous coupon payments in the CDS-based setting still does not significantly reduce bias. Although the averages are slightly higher than in the literal settings, the difference is minimal. These results indicate that modifications to default boundaries and

maturity and continuous coupon payments—tend to lead to weaker performance, particularly in hedging, relative to the Merton model.

coupon payments have limited effects in addressing these biases.

[Insert Figure 7 here]

Although the rank correlations are lower than in the Merton model, convergence signals might be stronger, as adding these two factors may enhance relative valuations cross-sectionally. Table 8 demonstrates the performance of bond portfolios with cross-sectional debiasing and combined cross-sectional and time-series debiasing in both literal and CDS-based settings. Panel A shows that, for cross-sectional debiasing alone, the literal setting of the Leland-Toft model exhibits the worst performance, with information ratios of 1.55 for d at 0.01, 1.55 for d at 0.1, and 1.52 for d at 0.3, compared to the Merton model's 1.89.

For combined debiasing, it also shows worse performance, with information ratios of 2.18 for d at 0.01, 2.1 for d at 0.1, and 2.1 for d at 0.3, compared to the Merton model's 2.52. The portfolio standard deviations are also higher than those of the Merton model, indicating that the literal setting of Leland-Toft does not provide a better valuation. Instead, these adjustments disrupt the relative ranking convergence signals. Consequently, the delta-hedging performances in Panel B worsen as the default barriers increase.

Panel C shows that, for cross-sectional debiasing alone, the CDS-based Leland-Toft model with R at 0.7 exhibits worse performance, with information ratios of 1.64 for d at 0.7, 1.76 for d at 0.8, and 1.87 for d at 0.9, compared to the Merton model's 1.99 with R at 0.7. Notably, their portfolio standard deviations are quite close, with the Merton model having 0.45 and 0.48 for d at 0.7, 0.45 for d at 0.8, and 0.40 for d at 0.9.

For combined debiasing, it also shows worse performance, with information ratios of 2.24 for d at 0.7, 2.39 for d at 0.8, and 2.38 for d at 0.9, compared to the Merton model's 3.01 with R at 0.7. The portfolio standard deviations are also higher than those of the Merton model, indicating that the CDS-based Leland-Toft model still does not provide a better valuation after time-series debiasing. The delta-hedging performances in Panel D seem to effectively reduce portfolio risk; however, the risks remain higher than those of the Merton model.

[Insert Table 8 here]

Although structural models incorporate default boundaries and coupon payment assumptions to capture corporate policies, they are limited in explaining relative valuation and improving convergence signals. These models do not have significant abilities to address bias and still require time-series debiasing. Therefore, bias persists in these models, and the Merton model appears to perform the best compared to the Leland-Toft models with different default barriers, regardless of whether literal or CDS-based settings are used.

6 Conclusion

This paper evaluates the empirical investment performance of corporate bond pricing structural models. Although these models provide a theoretically sound framework for valuing corporate bonds and estimating default probabilities, credit spreads, and hedging ratios, their practical application, especially for hedging, is limited by persistent biases. These biases, driven by the financial constraints of the firms and the assumptions of the model, lead to a misclassification of credit risk, particularly among investment-grade firms.

To address these limitations, I propose two debiasing approaches using cross-sectional and time-series adjustments. Cross-sectional debiasing corrects for bias and scale issues in relative rankings, while time-series debiasing uses six-month rolling regressions to align model-implied spreads with actual credit spreads. In essence, this approach integrates investors' risk perceptions in corporate bond markets to refine model-implied spread predictions, producing more accurate market-level credit risk rankings. Together, these adjustments improve the rank correlation between actual and model-implied credit spreads to an average of 89%, suggesting that structural models remain effective in capturing relative valuations between firms and in identifying potential mispricing.

Beyond model biases, corporate bond markets are also influenced by structural risks. Effective portfolio construction for capital structure arbitrage requires three key criteria: (1) unpredictable pricing errors, (2) consistent risk exposure across portfolio quintiles,

and (3) an effective hedging mechanism using stock. The first two criteria aim to neutralize structural risk exposures, while the third functions as a delta-hedging strategy against underlying exposures when bond prices are sensitive to stock price movements. The results suggest that delta-hedging the underlying risk alone is not sufficient to improve the information ratio. However, it serves as a first step toward more effective protection against structural risks. By combining delta-hedging with long-short positions to neutralize structural risks, the overall hedging strategy becomes more effective, resulting in a significantly higher information ratio. Thus, these three criteria are inherently interconnected.

Using only cross-sectional debiasing, the pricing error reduces portfolio risk by 16% in a long-short delta-hedged bond portfolio. When based on both cross-sectional and time-series debiasing, the pricing error more effectively neutralizes structural risks for most of the sample period, achieving a 46% risk reduction. Compared to the standalone long and short positions, the strategy cuts risk by over 70%. It achieves an annualized information ratio of 2.95 approaching a low-risk arbitrage over a 20-year horizon. From September 2002 to December 2021, the hedged portfolio experienced negative returns in only 25 months, and those losses were relatively small, underscoring the strategy's stability despite some residual risks. Across two- to eight-month holding horizons, delta hedging improves the information ratio up to seven months, while valuations converge at both market and firm levels, indicating temporary stock–bond mispricing.

Using the twisted CDS-based model setup further improves the hedging performance, reducing the risk by approximately 60% compared to an unhedged portfolio. This demonstrates the model's ability to extract cleaner signals to measure sensitivity between stock and bond prices through the two debiasing techniques, whereas relying solely on the cross-sectional approach can amplify flawed hedging inferences. The Leland-Toft model, in its literal setting, leads to ineffective hedging strategies. Although the CDS-based Leland-Toft model with varying default boundaries shows solid performance, the Merton model consistently outperforms it. This suggests that applying the two debiasing techniques reduces

the differences among structural models.

This paper aims to make three key contributions: (1) It reinterprets the credit spread puzzle from an investment perspective, showing that model misspecification primarily affects hedging performance rather than valuation alone. (2) The proposed cross-sectional and time-series debiasing methods offer simple yet powerful tools for improving the empirical accuracy of structural credit models. (3) The results provide actionable guidance for asset managers on constructing bond portfolios that hedge structural and underlying risks while maintaining high information ratios. Together, these contributions bridge the gap between structural credit theory and empirical portfolio management. The findings provide a new framework for evaluating the real-world utility of credit risk models.

Future research could enhance these debiasing methods by constructing long-short portfolios with more precise error control to eliminate additional residual structural risks. It could also focus on developing new structural models that incorporate richer accounting information to better capture fundamental values, thereby strengthening convergence or divergence signals at both the market and firm levels after debiasing.

References

Abreu, D. and M. K. Brunnermeier (2002). Synchronization Risk and Delayed Arbitrage. *Journal of Financial Economics, 66*(2-3), pp. 341–360.

Acharya, V. V., Y. Amihud, and S. T. Bharath (2013). Liquidity Risk of Corporate Bond Returns: Conditional Approach. *Journal of Financial Economics, 110*(2), pp. 358–386.

Almeida, C., J. Fan, G. Freire, and F. Tang (2023). Can a Machine Correct Option Pricing Models? *Journal of Business Economic Statistics, 41*(3), pp. 995–1009.

Amihud, Y. and H. Mendelson (1986). Asset Pricing and the Bid–Ask Spread. *Journal of Financial Economics, 17*(2), pp. 223–249.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang (2006). The Cross-Section of Volatility and Expected Returns. *The Journal of Finance, 61*(1), pp. 259–299.

Antill, S. (2022). Do the Right Firms Survive Bankruptcy? *Journal of Financial Economics, 144*(2), pp. 523–546.

Antill, S. and S. R. Grenadier (2019). Optimal Capital Structure and Bankruptcy Choice: Dynamic Bargaining vs Liquidation. *Journal of Financial Economics, 133*(1), pp. 198–224.

Asquith, P., A. S. Au, T. Covert, and P. A. Pathak (2013). The Market for Borrowing Corporate Bonds. *Journal of Financial Economics, 107*(1), pp. 155–182.

Atmaz, A., S. Basak, and F. Ruan (2024). Dynamic Equilibrium with Costly Short-Selling and Lending Market. *The Review of Financial Studies, 37*(2), pp. 444–506.

Bai, J. and P. Collin-Dufresne (2019). The CDS-bond Basis. *Financial Management, 48*(2), pp. 417–439.

Bai, J., R. S. Goldstein, and F. Yang (2020). Is the Credit Spread Puzzle a Myth? *Journal of Financial Economics, 137*(2), pp. 297–319.

Bai, J. and L. Wu (2016). Anchoring Credit Default Swap Spreads to Firm Fundamentals. *The Journal of Financial and Quantitative Analysis, 51*(5), pp. 1521–1543.

Bakshi, G., X. Gao, and Z. Zhong (2022). Decoding Default Risk: A Review of Modeling Approaches, Findings, and Estimation Methods. *Annual Review of Financial Economic, 14*, pp. 391–413.

Bao, J. and K. Hou (2017). De Facto Seniority, Credit Risk, and Corporate Bond Prices. *The Review of Financial Studies, 30*(11), pp. 4038–4080.

Bao, J., J. Pan, and J. Wang (2011). The Illiquidity of Corporate Bonds. *The Journal of Finance, 66*(3), pp. 911–946.

Beber, A. and M. Pagano (2013). Short-Selling Bans Around the World: Evidence from the 2007–09 Crisis. *The Journal of Finance, 68*(1), pp. 343–381.

Bessembinder, H., S. E. Jacobsen, W. F. Maxwell, and K. Venkataraman (2018). Capital Commitment and Illiquidity in Corporate Bonds. *The Journal of Finance, 73*(4), pp. 1615–1661.

Bessembinder, H., K. M. Kahle, W. F. Maxwell, and D. Xu (2009). Measuring Abnormal Bond Performance. *The Review of Financial Studies, 22*(10), pp. 4219–4258.

Bharath, S. T. and T. Shumway (2008). Forecasting Default with the Merton Distance to Default Model. *The Review of Financial Studies, 21*(3), pp. 1339–1369.

Black, F. and J. C. Cox (1976). Valuing Corporate Securities: Some Effects of Bond Indenture Provisions. *The Journal of Finance, 31*(2), pp. 351–367.

Blanco, R., S. Brennan, and I. W. Marsh (2005). An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps. *The Journal of Finance, 60*(5), pp. 2255–2281.

Bris, A., W. Goetzmann, and N. Zhu (2007). Efficiency and the Bear: Short Sales and Markets Around the World. *The Journal of Finance, 62*(3), pp. 1029–1079.

Brunnermeier, M. K. and L. H. Pedersen (2009). Market Liquidity and Funding Liquidity. *The Review of Financial Studies, 22*(6), pp. 2201–2238.

Brunnermeier, M. K. and Y. Sannikov (2014). A Macroeconomic Model with a Financial Sector. *The American Economic Review, 104*(2), pp. 379–421.

Cao, J. and B. Han (2013). Cross Section of Option Returns and Idiosyncratic Stock Volatility. *Journal of Financial Economics, 108*(1), pp. 231–249.

Chen, H., Z. Chen, and J. Li (2023). The Debt-Equity Spread. *Working Paper*.

Chen, H., R. Cui, Z. He, and K. Milbradt (2018). Quantifying Liquidity and Default Risks of Corporate Bonds over the Business Cycle. *The Review of Financial Studies, 31*(3), pp. 852–897.

Chen, L., D. A. Lesmond, and J. Wei (2007). Corporate Yield Spreads and Bond Liquidity. *The Journal of Finance, 62*(1), pp. 119–149.

Chen, Z. and P. J. Knez (1995). Measurement of Market Integration and Arbitrage. *The Review of Financial Studies, 8*(2), pp. 287–325.

Choi, J. and Y. Kim (2018). Anomalies and Market (Dis)Integration. *Journal of Monetary Economics, 100*, pp. 16–34.

Choi, J. and M. Kronlund (2018). Reaching for Yield in Corporate Bond Mutual Funds. *The Review of Financial Studies, 31*(5), pp. 1930–1965.

Choi, J. and M. Richardson (2016). The Volatility of a Firm's Assets and the Leverage Effect. *Journal of Financial Economics, 121*(2), pp. 254–277.

Chordia, T., A. Goyal, Y. Nozawa, A. Subrahmanyam, and Q. Tong (2017). Are Capital Market Anomalies Common to Equity and Corporate Bond Markets? An Empirical Investigation. *The Journal of Financial and Quantitative, Analysis, 52*(4), pp. 1301–1342.

Collin-Dufresn, P., R. S. Goldstein, and J. S. Martin (2001). The Determinants of Credit Spread Changes. *The Journal of Finance, 56*(6), pp. 2177–2207.

Collin-Dufresne, P., B. Junge, and A. B. Trolle (2024). How Integrated are Credit and Equity Markets? Evidence from Index Options. *The Journal of Finance, 79*(2), pp. 949–992.

Covitz, D. and C. Downing (2007). Liquidity or Credit Risk? The Determinants of Very Short-Term Corporate Yield Spreads. *The Journal of Finance, 62*(5), pp. 2303–2328.

Crouzet, N. and F. Tourre (2021). Can the Cure Kill the Patient? Corporate Credit Interventions and Debt Overhang. *Working Paper*.

Dick-Nielsen, J., P. Feldhütter, and D. Lando (2012). Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis. *Journal of Financial Economics, 103*(3), pp. 471–492.

Dickerson, A., P. Mueller, and C. Robotti (2023). Priced Risk in Corporate Bonds. *Journal of Financial Economics, 150*(2).

Doshi, H., K. Jacobs, P. Kumar, and R. Rabinovitch (2019). Leverage and the Cross-Section of Equity Returns. *The Journal of Finance, 74*(3), pp. 1431–1471.

Du, D., R. Elkamhi, and J. Ericsson (2019). Time-Varying Asset Volatility and the Credit Spread Puzzle. *The Journal of Finance, 74*(4), pp. 1841–1885.

Duarte, J., F. A. Longstaff, and F. Yu (2007). Risk and Return in Fixed-Income Arbitrage: Nickels in Front of a Steamroller? *The Review of Financial Studies, 20*(3), pp. 769–811.

Duffee, G. R. (1998). The Relation Between Treasury Yields and Corporate Bond Yield Spreads. *The Journal of Finance, 53*(6), pp. 2225–2241.

Duffie, D. (2010). Presidential Address: Asset Price Dynamics with Slow-Moving Capital. *The Journal of Finance, 65*(4), pp. 1237–1267.

Duffie, D., N. Garleanu, and L. Pedersen (2002). Securities Lending, Shorting, and Pricing. *Journal of Financial Economics, 66*(2-3), pp. 307–339.

Eom, Y. H., J. Helwege, and J.-Z. Huang (2004). Structural Models of Corporate Bond Pricing: An Empirical Analysis. *The Review of Financial Studies, 17*(2), pp. 499–544.

Ericsson, J., K. Jacobs, and R. Oviedo (2009). The Determinants of Credit Default Swap Premia. *The Journal of Financial and Quantitative Analysis, 44*(1), pp. 109–132.

Fama, E. F. and K. R. French (1993). Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics, 33*(1), pp. 3–56.

Fama, E. F. and K. R. French (2005). Financing Decisions: Who Issues Stock? *Journal of Financial Economics, 76*(3), pp. 549–582.

Faulkender, M. W. (2005). Hedging or Market Timing? Selecting the Interest Rate Exposure of Corporate Debt. *The Journal of Finance, 60*(2), pp. 931–962.

Fazzari, S., R. G. Hubbard, and B. C. Petersen (1988). Financing Constraints and Corporate Investment. *Brookings Papers on Economic Activity, 19*(1), pp. 141–206.

Feldhütter, P. and S. Schaefer (2023). Debt Dynamics and Credit Risk. *Journal of Financial Economics, 149*(3), pp. 497–535.

Feldhütter, P. and S. M. Schaefer (2018). The Myth of the Credit Spread Puzzle. *The Review of Financial Studies, 31*(8), pp. 2897–2942.

Friewald, Jankowitsch, and Subrahmanyam (2012). Illiquidity or Credit Deterioration: A Study of Liquidity in the US Corporate Bond Market During Financial Crises. *Journal of Financial Economics, 105*(1), pp. 18–36.

Friewald, N., C. Wagner, and J. Zechner (2014). The Cross-Section of Credit Risk Premia and Equity Returns. *The Journal of Finance, 69*(6), pp. 2419–2469.

Geske, R. (1977). The Valuation of Corporate Liabilities as Compound Options. *The Journal of Financial and Quantitative Analysis, 12*(4), pp. 541–552.

Gilchrist, S. and E. Zakrajšek (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review, 102*(4), pp. 1692–1720.

Goldberg, J. E. and Y. Nozawa (2021). Liquidity Supply in the Corporate Bond Market. *The Journal of Finance, 76*(2), pp. 755–796.

Goldstein, M. A. and E. S. Hotchkiss (2020). Providing Liquidity in an Illiquid Market: Dealer Behavior in US Corporate Bonds. *Journal of Financial Economics, 135*(1), pp. 16–40.

Goyal, A., Y. Nozawa, and Y. Qiu (2025). Passive Ownership and Corporate Bond Lending. *Working Paper*.

Gromb, D. and D. Vayanos (2010). Limits of Arbitrage. *Annual Review of Financial Economics, 2*, pp. 251–275.

Grullon, G., S. Michenaud, and J. Weston (2015). The Real Effects of Short-Selling Constraints. *The Review of Financial Studies, 28*(6), pp. 1737–1767.

He, Z. and A. Krishnamurthy (2013). Intermediary Asset Pricing. *The American Economic Review, 103*(2), pp. 732–770.

He, Z. and K. Milbradt (2014). Endogenous Liquidity and Defaultable Bonds. *Econometrica, 82*(4), pp. 1443–1508.

He, Z. and W. Xiong (2012). Rollover Risk and Credit Risk. *The Journal of Finance, 67*(2), pp. 391–430.

Hendershott, T., R. Kozhan, and V. Raman (2020). Short Selling and Price Discovery in Corporate Bonds. *The Journal of Financial and Quantitative Analysis, 55*(1), pp. 77–115.

Hong, H. and J. C. Stein (2003). Differences of Opinion, Short-Sales Constraints, and Market Crashes. *The Review of Financial Studies, 16*(2), pp. 487–525.

Huang, J.-Z. and M. Huang (2012). How Much of the Corporate-Treasury Yield Spread Is Due to Credit Risk? *The Review of Asset Pricing Studies, 2*(2), pp. 153–202.

Huang, J.-Z., B. Liu, and Z. Shi (2023). Determinants of Short-Term Corporate Yield Spreads: Evidence from the Commercial Paper Market. *Review of Finance, 27*(2), pp. 539–579.

Huang, J.-Z., Y. Nozawa, and Z. Shi (2025). The Global Credit Spread Puzzle. *The Journal of Finance, 80*(1), pp. 101–162.

Huang, J.-Z. and Z. Shi (2021). What Do We Know About Corporate Bond Returns? *Annual Review of Financial Economics, 13*, pp. 363–399.

Huang, J.-Z., Z. Shi, and H. Zhou (2020). Specification Analysis of Structural Credit Risk Models. *Review of Finance, 24*(1), pp. 45–98.

ICMA (2019). What Is "Short Selling" and What Is the Role of Repo? *Frequently Asked Questions on Repo, Question 30*, International Capital Market Association.

Jegadeesh, N. and S. Titman (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance, 48*(1), pp. 65–91.

Jones, E. P., S. P. Mason, and E. Rosenfeld (1984). Contingent Claims Analysis of Corporate Capital Structures: an Empirical Investigation. *The Journal of Finance, 39*(3), pp. 611–625.

Kapadia, N. and X. Pu (2012). Limited Arbitrage between Equity and Credit Markets. *Journal of Financial Economics, 105*(3), pp. 542–564.

Kaplan, S. N. and L. Zingales (1997). Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints? *The Quarterly Journal of Economics, 112*(1), pp. 169–215.

Kelly, B. T., D. Palhares, and S. Pruitt (2023). Modeling Corporate Bond Returns. *The Journal of Finance, 78*(4), pp. 1967–2008.

Kuong, J. C.-F., J. O'Donovan, and J. Zhang (2024). Monetary Policy and Fragility in Corporate Bond Mutual Funds. *Journal of Financial Economics, 161*.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica, 53*(6), pp. 1315–1335.

Lagos, R. and G. Rocheteau (2009). Liquidity in Asset Markets with Search Frictions. *Econometrica, 77*(2), pp. 403–426.

Leland, H. E. (1994). Corporate Debt Value, Bond Covenants, and Optimal Capital Structure. *The Journal of Finance, 49*(4), pp. 1213–1252.

Leland, H. E. and K. B. Toft (1996). Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. *The Journal of Finance, 51*(3), pp. 987–1019.

Li, J. and M. C. Ringgenberg (2024). Do Interventions in Secondary Markets Affect the Economy? *Working Paper*.

Li, J. S., M. O'Hara, A. C. Rapp, and X. A. Zhou (2025). Bond Market Illiquidity: Is Portfolio Trading the Solution? *Working Paper*.

Li, Z. (2024). Long Rates, Life Insurers, and Credit Spreads. *Working Paper*.

Lin, H., J. Wang, and C. Wu (2011). Liquidity Risk and Expected Corporate Bond Returns. *Journal of Financial Economics, 99*(3), pp. 628–650.

Longstaff, F. A., S. Mithal, and E. Neis (2005). Corporate Yield Spreads: Default Risk or Liquidity? New Evidence from the Credit Default Swap Market. *The Journal of Finance, 60*(5), pp. 2213–2253.

Lowenstein, R. (2000). When Genius Failed: The Rise and Fall of Long-Term Capital Management. *Random House*.

Ma, Y. (2019). Nonfinancial Firms as Cross-Market Arbitrageurs. *The Journal of Finance, 74*(6), pp. 3041–3087.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance, 7*(1), pp. 77–91.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance, 29*(2), pp. 449–470.

Miller, E. M. (1977). Risk, Uncertainty, and Divergence of Opinion. *The Journal of Finance, 32*(4), pp. 1151–1168.

Modigliani, F. and M. H. Miller (1958). The Cost of Capital, Corporation Finance and the Theory of Investment. *The American Economic Review, 48*(3), pp. 261–297.

Rajan, R. G. and L. Zingales (1998). Financial Dependence and Growth. *The American Economic Review, 88*(3), pp. 559–586.

Ross, S. A. (1976). The Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory, 13*(3), pp. 341–360.

Sandulescu, M. (2022). How Integrated Are Corporate Bond and Stock Markets? *Working Paper*.

Schaefer, S. M. and I. A. Strebulaev (2008). Structural Models of Credit Risk are Useful: Evidence from Hedge Ratios on Corporate Bonds. *Journal of Financial Economics, 90*(1), pp. 1–19.

Shleifer, A. and R. W. Vishny (1997). The Limits of Arbitrage. *The Journal of Finance*, *52*(1), pp. 35–55.

Sufi, A. (2009). Bank Lines of Credit in Corporate Finance: An Empirical Analysis. *The Review of Financial Studies*, *22*(3), pp. 1057—1088.

Tian, M. and L. Wu (2023). Limits of Arbitrage and Primary Risk-Taking in Derivative Securities. *The Review of Asset Pricing Studies*, *13*(3), pp. 405–439.

van Binsbergen, J. H., Y. Nozawa, and M. Schwert (2025). Duration-Based Valuation of Corporate Bonds. *The Review of Financial Studies*, *38*(1), pp. 158—-191.

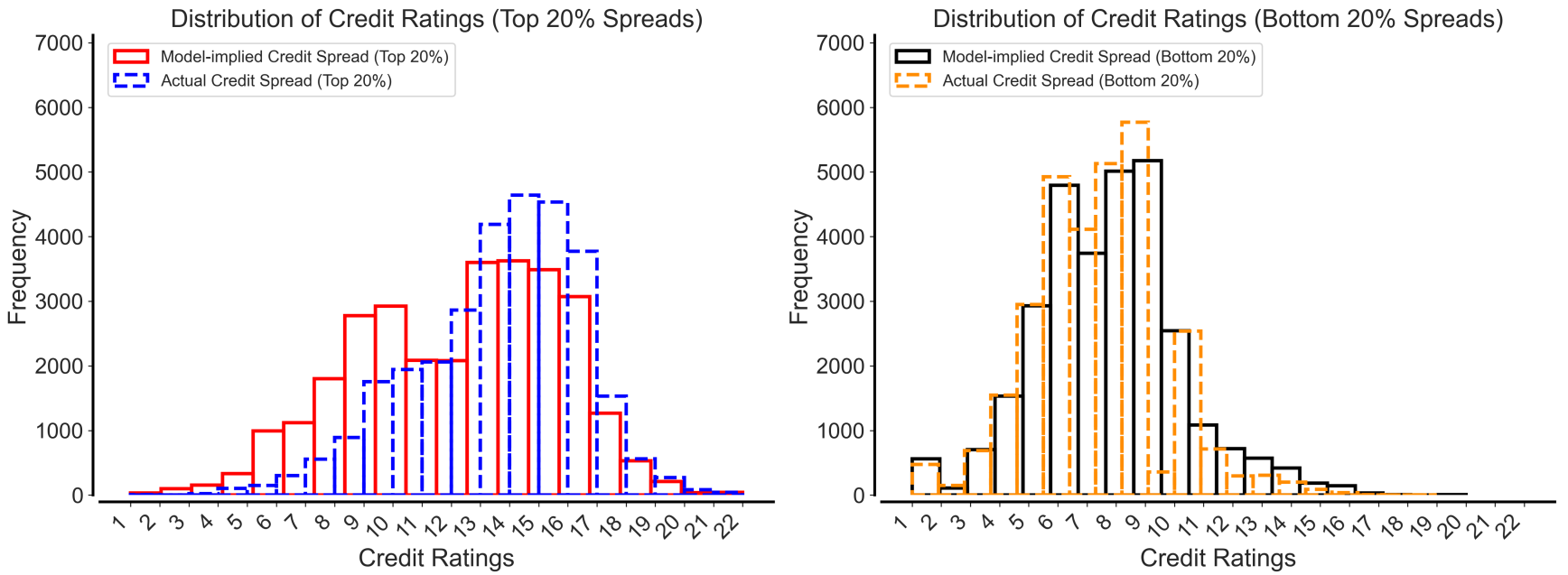
Vayanos, D. and P.-O. Weill (2008). A Search-Based Theory of the On-the-Run Phenomenon. *The Journal of Finance*, *63*(3), pp. 1361–1398.

Wu, L. and H. Zaman (2024). Finding Value in the U.S. Corporate Bond Market. *Working Paper*.

Xiong, W. (2001). Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets. *Journal of Financial Economics*, *62*(2), pp. 247–292.

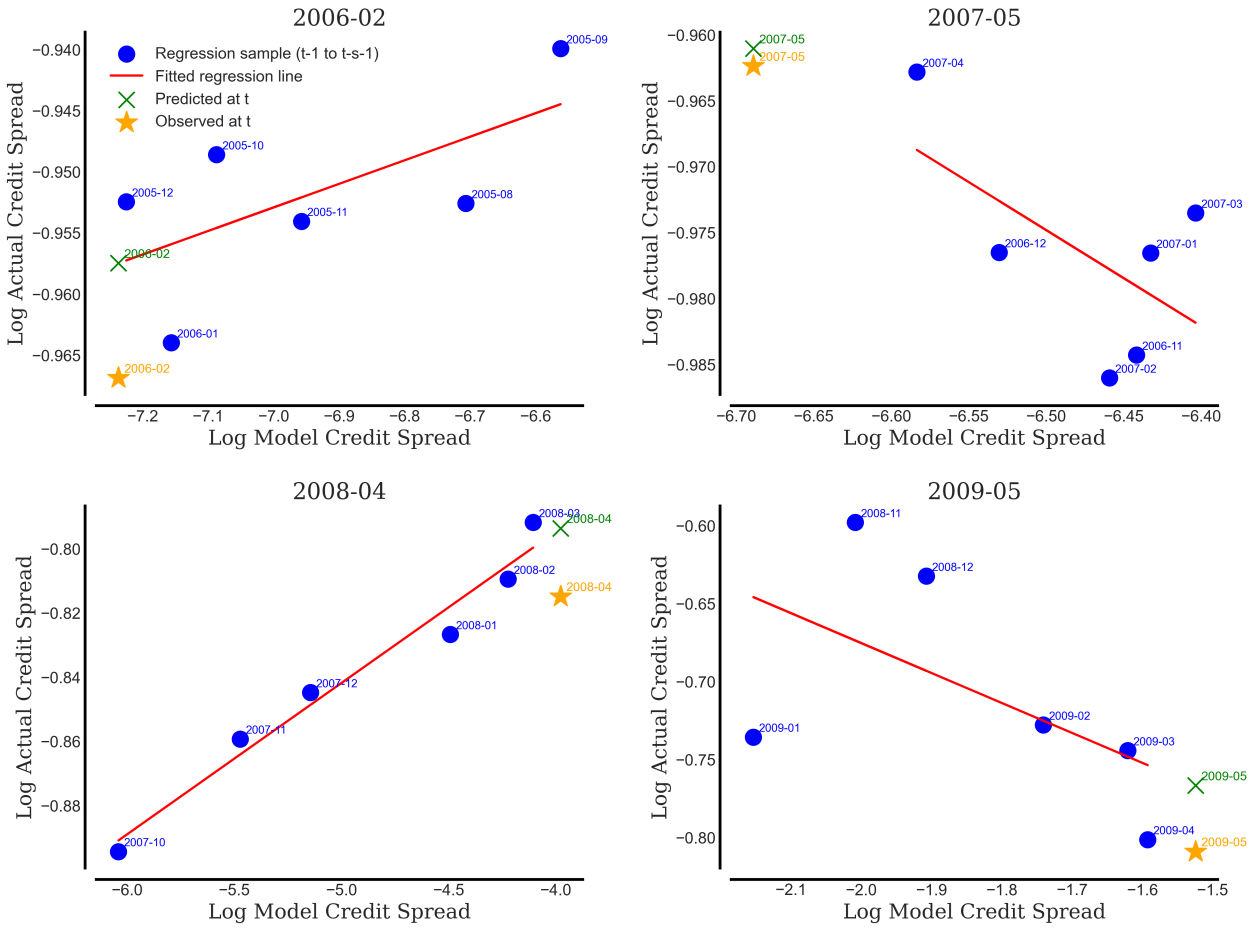
Yu, F. (2006). How Profitable Is Capital Structure Arbitrage? *Financial Analysts Journal*, *62*(5), pp. 47–62.

Figure 1: **Distribution of Credit Ratings: Top 20% vs. Bottom 20% Credit Spreads**



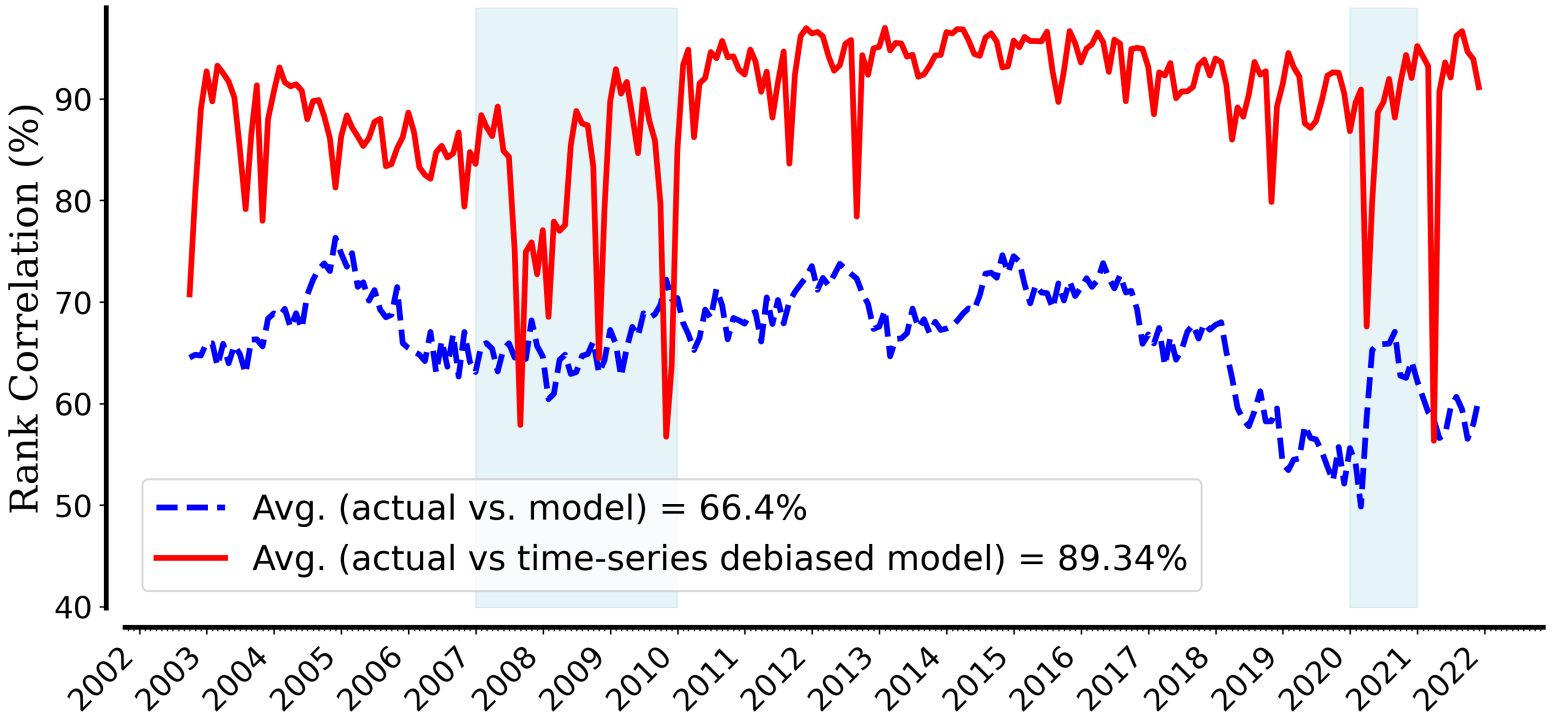
This figure shows the histograms for the top 20% and bottom 20% of both actual and model-implied credit spreads (based on the literal setting of the Merton model) as defined in Equation (3). The estimation process for asset value and its volatility follows the formulations in Equations (1) and (2). The model inputs are outlined in Section (4.2). The x-axis represents the S&P Bond Issue Credit Ratings, ranging from 1 to 22, as provided in the WRDS Bond Returns dataset. A rating of 1 corresponds to the highest credit quality (AAA), while 22 represents the lowest rating (D). Ratings above 11 are classified as investment-grade, while ratings of 11 or below are categorized as high-yield. The y-axis shows the frequency, or the number of data points, within each rating category, offering insights into the distribution of credit spreads across different rating classes.

Figure 2: **Forecasting Trends with Time-Series Debiasing**



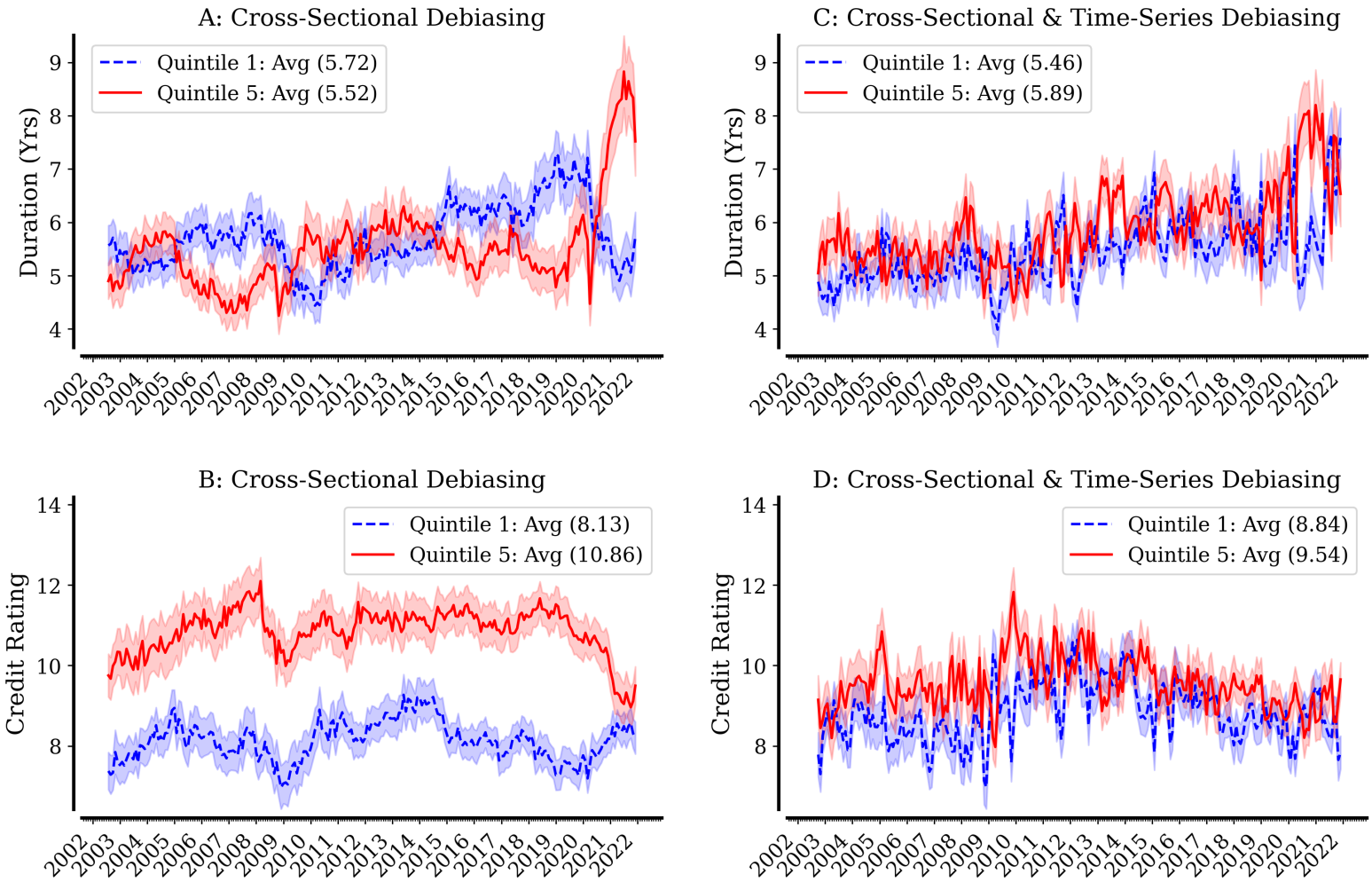
This figure illustrates the trend implications of time-series debiasing, using Centex Corporation (GVKEY 2845) as an example. The graph shows trends for four specific months: February 2006, May 2007, April 2008, and May 2009. For each of these months, a time-series regression is conducted based on data from the prior six months. This regression estimates the intercept and slope, which are used to model-implied credit spread for the following month. The predicted credit spread from the model is then compared with the actual observed spread for that month to assess the model's accuracy. All regressions are carried out in log space.

Figure 3: **Rank Correlation Analysis**



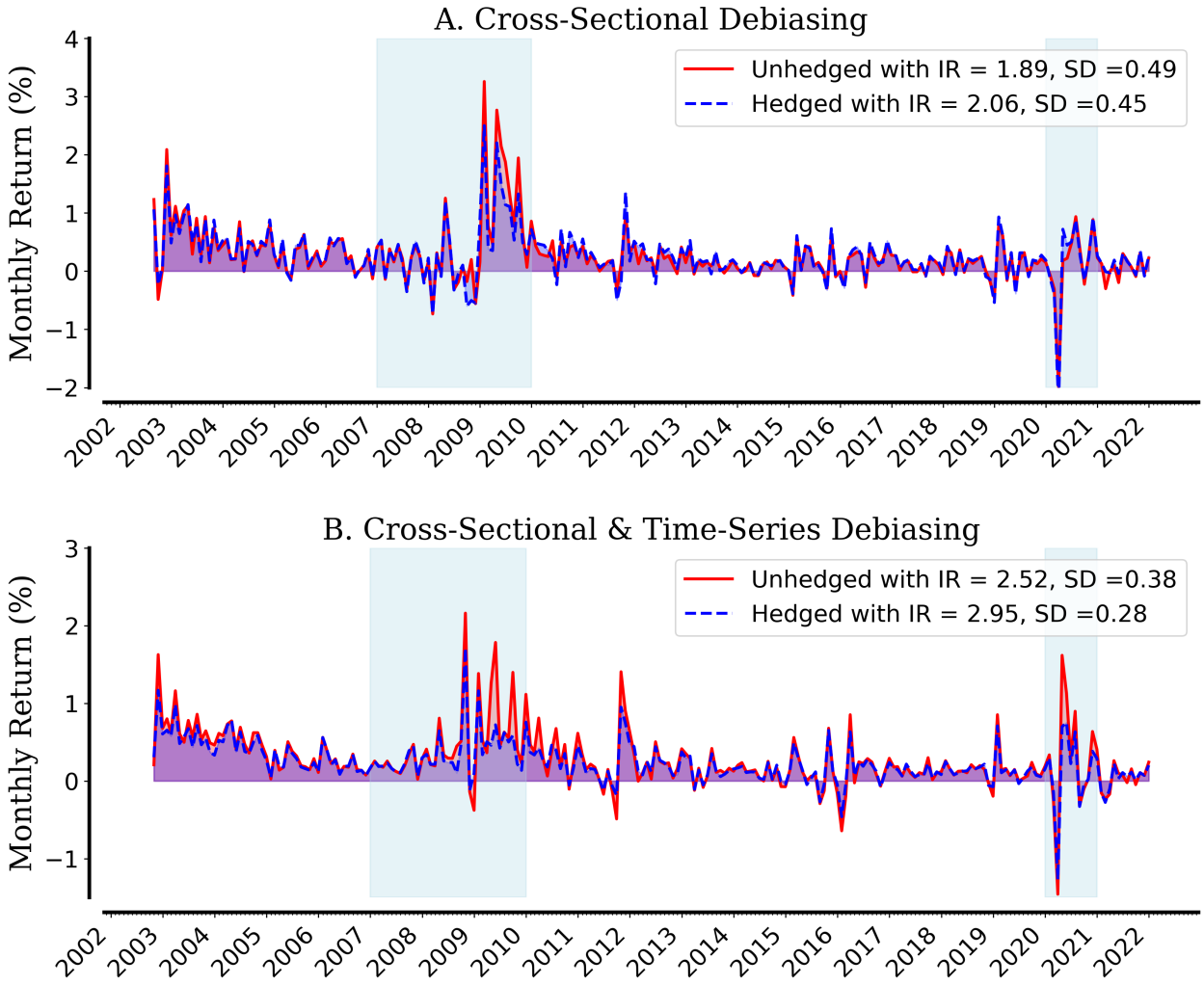
This figure presents the rank correlations between the actual credit spreads and two different sets of model spreads: the literal Merton model setting and its time-series debiased model. The blue dashed line represents the rank correlation between the actual spreads and the original model spreads, which is 66.4% on average. The red solid line shows the rank correlation between the actual spreads and the time-series debiased model spreads, which is 89.34%. The time-series debiasing process enhances the model by adjusting firms, especially investment-grade firms, to more accurate positions for comparing their credit risk, resulting in a higher correlation with the actual spreads. The data spans from September 2002 to November 2021.

Figure 4: **Comparison of Debiasing Methods on Duration and Credit Rating**



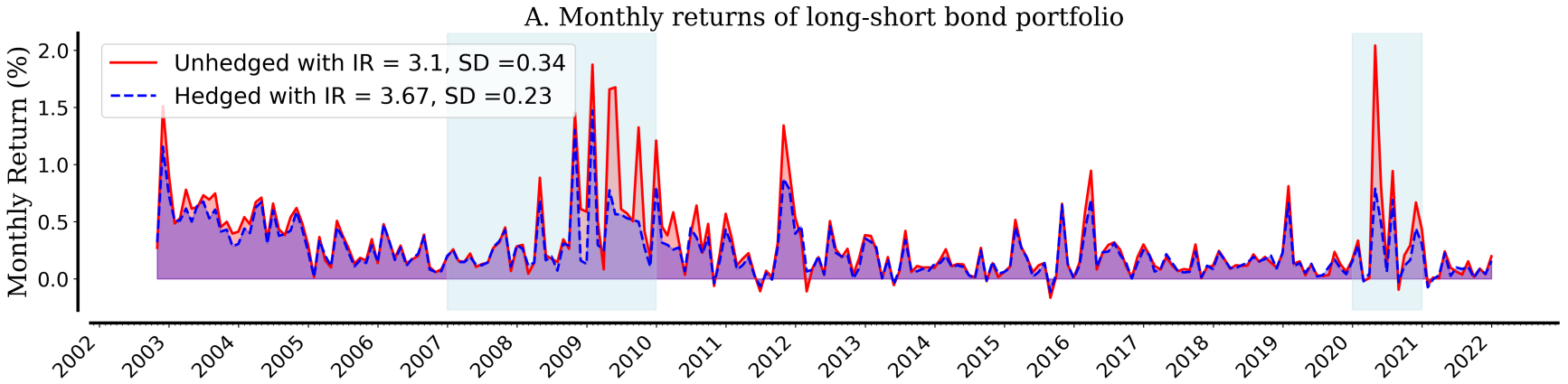
This figure presents the average values and 95% confidence intervals for duration and credit rating in quintiles 1 (blue dashed line) and 5 (red solid line) over time. In Panels A and B, the error under cross-sectional debiasing is defined as the rank difference between the actual credit spreads and the model-implied credit spreads. In Panels C and D, the error under combined cross-sectional and time-series debiasing is defined as the rank difference between the actual and model-implied spreads after correcting for time-series bias.

Figure 5: **Unlevered Return Analysis of Long-Short Bond Portfolio**



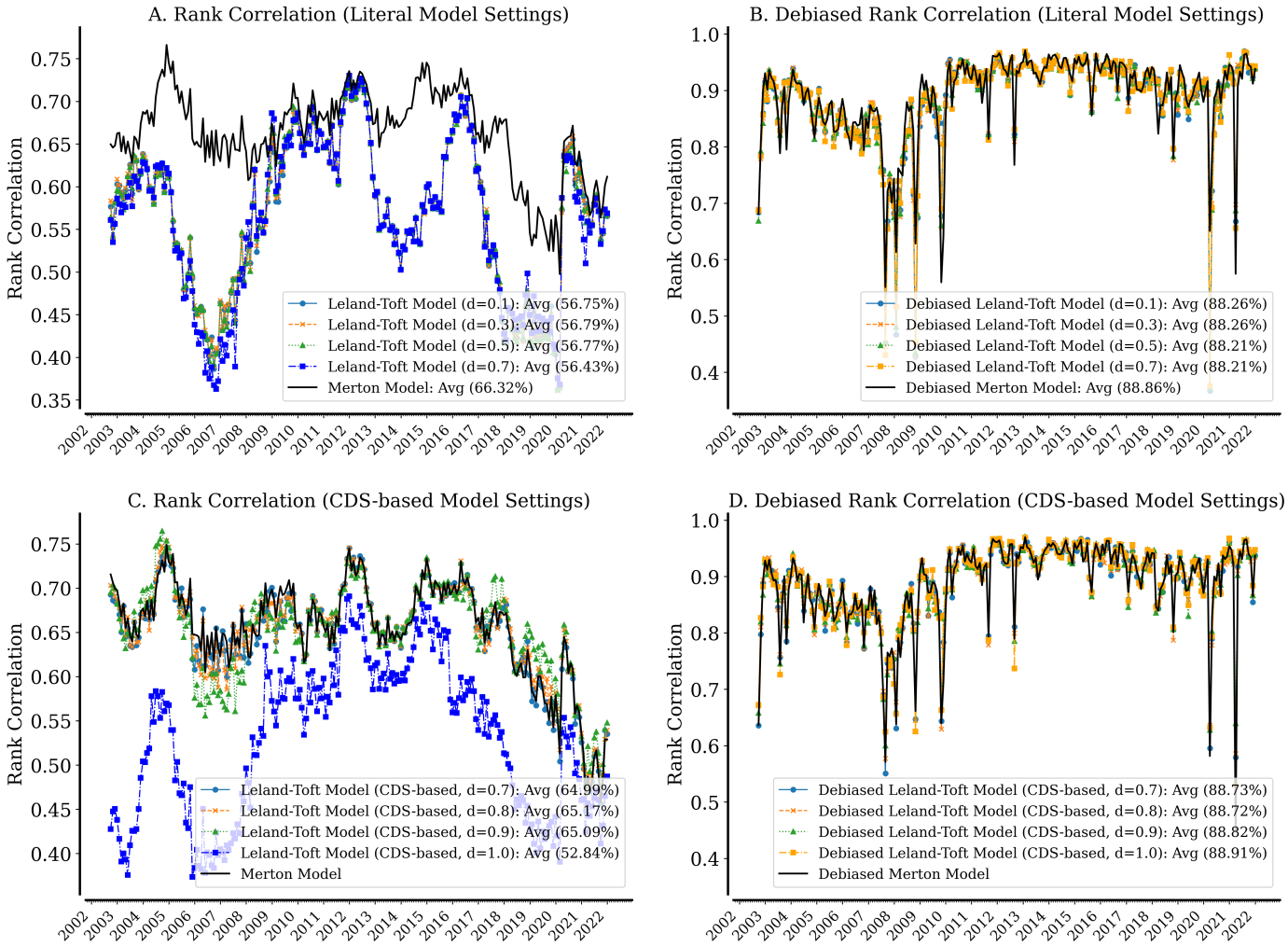
This figure illustrates the performance of the long-short bond portfolio under two settings. Panel A displays the monthly unlevered returns of the unhedged (red solid line) and delta-hedged (blue dashed line) portfolios using only cross-sectional debiasing. Panel B shows the corresponding returns when both cross-sectional and time-series debiasing are applied.

Figure 6: **Holding Return Analysis of Long-Short Bond Portfolio**



This figure illustrates the performance of the long-short bond portfolio for extending the holding period to the next month if monthly returns are negative, for a maximum of six months. The monthly unlevered returns for the unhedged (red solid line) and delta-hedged (blue dashed line) portfolios.

Figure 7: **Rank Correlation Analysis: Merton vs. Leland-Toft Models with Varying Default Barriers**



This figure compares the rank correlations between actual credit spreads and those predicted by the Merton and Leland-Toft models for different default barrier settings. For the literal setting, $d = 0.1, 0.3, 0.5, 0.7$, and for the CDS-based setting ($R = 0.7$), $d = 0.7, 0.8, 0.9, 1.0$. Panel A reports the rank correlation between actual and model-implied credit spreads for the literal models, while Panel B shows the results after applying the time-series debiasing. Panel C presents the rank correlation for the CDS-based models, and Panel D shows the results with the time-series debiased CDS-based models.

Table 1: **Summary Statistics**

	Mean	SD	Percentile Values					Tails	
			1	25	50	75	99	Skewness	Kurtosis
					A. Data				
Excess Stock Returns (%)	14.17	126.38	-306.13	-46.73	12.43	70.87	374.92	2.54	75.96
Excess Bond Returns (%)	5.03	33.82	-80.3	-5.22	4.01	15.1	96.74	3.32	120.58
Stock Trading Price	119.69	2401.2	5.93	23.17	40.05	65.67	423.33	42.76	1903.56
Bond Trading Price	104.28	8.17	77.1	100.6	104.55	108.68	122.6	-0.29	42.8
YTM (%)	5.06	2.77	1.21	3.28	4.76	6.28	13.28	5.87	122.6
1-Month Treasury Rate (%)	1.15	1.47	0.0	0.05	0.26	1.73	5.18	1.39	0.9
Actual Credit Spread (%)	3.92	2.85	0.35	2.18	3.34	5.1	12.75	5.82	114.32
Coupon (%)	5.96	1.76	2.5	4.67	5.92	7.08	10.5	0.33	0.16
Duration	5.58	2.3	1.33	4.02	5.26	6.8	12.29	0.86	1.31
Market Value of Debt ($M)	38,129	167,989	292	2,173	5,960	18,953	734,140	11	147
Market Value of Asset ($M)	61,438	200,047	580	5,176	14,333	43,113	891,136	9	113
					B. Inputs for the Merton Model				
Total Debt ($M)	36,150	157,984	291	2,110	5,730	17,850	703,677	11	138
MarketCap ($M)	23,309	69,513	202	2,265	6,480	19,034	243,390	17	452
Debt-to-Equity	2.06	4.46	0.1	0.47	0.92	1.88	18.22	18.38	953.52
Bond Maturity (Years)	8.24	4.56	1.55	5.08	7.06	10.43	23.25	1.3	2.11
Stock Volatility (%)	34.18	18.92	13.06	21.9	29.37	40.66	105.56	2.85	18.48
					C. Estimations for the Merton Model				
Asset ($M)	55,937	179,108	497	4,772	13,438	39,940	790,897	9	113
Asset Volatility (%)	17.58	13.07	2.31	10.06	15.5	21.97	67.16	4.64	58.29
					D. Outputs for the Merton Model				
Distance to Default, d_1	2.34	1.28	0.07	1.55	2.15	2.88	6.60	2.42	40.52
Credit Spread [Literal] (%)	0.77	3.79	0.00	0.01	0.07	0.39	14.04	19.56	657.64
Hedge Ratio [Literal] (%)	5.31	9.40	0.00	0.20	1.57	6.03	47.40	3.43	15.49
Credit Spread [CDS-Based] (%)	0.74	1.91	0.00	0.02	0.17	0.64	8.89	9.06	154.63
Hedge Ratio [CDS-Based] (%)	8.45	25.2	0.00	0.14	1.13	5.11	138.26	6.02	46.99

This table summarizes the time-series averages of the cross-sectional distribution of the data, reporting the cross-sectional mean (Mean), standard deviation (SD), percentiles, skewness, and excess kurtosis. Panel A displays statistics for the firm-level data sample, including annualized excess stock returns, annualized excess bond returns, stock trading price, bond trading price, actual yield to maturity (YTM), one-month Treasury rate (from Fed), actual credit spread, coupon, and duration. The market value of debt ($\overline{B}$) is calculated by multiplying the weighted bond price by the total debt, while the market value of assets is the sum of the market value of debt and market capitalization. These two variables are used to evaluate capital structure arbitrage performance in Section 3.3. Panel B reports statistics for the structural model inputs, including total debt (quarterly total liabilities from Compustat), market capitalization (monthly market capitalization from CRSP, calculated as stock price multiplied by shares outstanding), the debt-to-market capitalization ratio (debt-to-equity), bond maturity, and the one-year annualized daily stock return volatility estimator. Panel C shows the estimated asset value and asset return volatility from Equations (1) and (2). Panel D presents statistics for the model outputs, including distance to default (d_1), the model-implied credit spread from Equation (3), the model-implied hedge ratio from Equation (5), the CDS-based credit spread from Equation (4) assuming a recovery rate of $R = 0.7$, and the CDS-based hedge ratio from Equation (6), also using $R = 0.7$. Dollar values ($M) are in millions.

Table 2: **Summary Influences of Abnormal Yields in Portfolio Construction**

	Actual Spread				Model Spread				Spread Error					
Quintile	Mean	Rp	SD	IR	Mean	Rp	SD	IR	Actual Mean	Model Mean	Mean Error	Rp	SD	IR
A. All Firms														
All	3.90	0.22	0.76	1.02	0.77	0.22	0.76	1.02	3.90	0.77	3.13	0.22	0.76	1.02
1	1.88	0.02	0.43	0.18	0.02	0.08	0.49	0.58	2.63	2.16	0.47	0.18	0.74	0.82
2	2.71	0.10	0.65	0.53	0.07	0.17	0.66	0.89	2.73	0.33	2.41	0.09	0.62	0.50
3	3.40	0.18	0.75	0.81	0.16	0.19	0.75	0.89	3.34	0.32	3.02	0.15	0.73	0.72
4	4.58	0.22	0.88	0.88	0.46	0.26	0.84	1.05	4.39	0.45	3.94	0.21	0.86	0.85
5	6.92	0.60	1.54	1.35	3.11	0.42	1.35	1.08	6.39	0.56	5.83	0.49	1.25	1.36
5-1		0.58	1.39	1.44		0.34	1.11	1.06				0.32	0.89	1.23
B. Investment-Grade Firms														
All	2.90	0.15	0.68	0.79	0.25	0.15	0.68	0.79	2.90	0.25	2.65	0.15	0.68	0.79
1	1.70	0.01	0.39	0.12	0.01	0.06	0.43	0.48	1.87	0.51	1.36	0.05	0.48	0.33
2	2.39	0.07	0.56	0.41	0.05	0.14	0.65	0.75	2.39	0.17	2.22	0.07	0.55	0.41
3	2.81	0.12	0.71	0.57	0.10	0.17	0.73	0.79	2.78	0.17	2.62	0.11	0.68	0.57
4	3.27	0.17	0.79	0.76	0.21	0.18	0.77	0.81	3.22	0.18	3.04	0.16	0.78	0.73
5	4.33	0.40	1.10	1.26	0.88	0.22	0.91	0.85	4.23	0.23	4.00	0.38	1.05	1.27
5-1		0.39	0.84	1.59		0.16	0.60	0.94				0.34	0.70	1.67
C. High-Yield Firms														
All	5.80	0.37	1.17	1.11	1.75	0.37	1.17	1.11	5.80	1.75	4.06	0.37	1.17	1.11
1	3.79	0.08	0.71	0.40	0.08	0.24	0.82	1.01	5.03	5.09	-0.06	0.40	1.44	0.95
2	4.83	0.17	0.83	0.71	0.31	0.31	0.99	1.09	4.98	1.47	3.51	0.24	0.97	0.86
3	5.44	0.28	0.95	1.00	0.74	0.29	1.09	0.92	5.25	0.85	4.40	0.23	0.95	0.83
4	6.20	0.35	1.33	0.91	1.64	0.39	1.28	1.05	5.78	0.60	5.18	0.31	1.01	1.05
5	8.75	0.99	2.46	1.40	5.94	0.64	2.02	1.10	7.96	0.68	7.28	0.69	1.83	1.30
5-1		0.91	2.08	1.52		0.40	1.51	0.93				0.29	1.11	0.92

This table consists of three columns: quintile bond portfolios based on (1) the actual credit spreads, (2) the model-implied credit spreads from the Merton model (Model spread), and (3) the difference between the actual and model-implied credit spreads (Spread error). For each portfolio, the table reports the average monthly excess returns over next month (R_p) in percentage, standard deviation (SD) in percentage, and annualized information ratio (IR). The final column presents the average values of the actual spread in percentage, model spread in percentage, and spread error in percentage across all quintiles, categorized by spread error.

Table 3: **Bias Correction and Scale Adjustment in Cross-Sectional Structural Model Analysis**

	Average			Unhedged			Hedged			Variance Reduction	Rank Correlation
Quintile	Actual	Model	Pricing Error	Rp	SD	IR	Rp	SD	IR	(%)	(%)
A. All Firms											
All				0.22	0.76	1.02	0.14	0.59	0.84	39.73	66.51
1	2.79	0.72	-1.39	0.07	0.63	0.41	0.02	0.50	0.12	37.01	
2	3.61	1.09	-0.43	0.14	0.74	0.67	0.07	0.58	0.43	38.57	
3	4.27	1.59	0.00	0.29	0.93	1.08	0.17	0.66	0.87	49.64	
4	4.13	0.33	0.44	0.27	0.82	1.14	0.18	0.64	0.98	39.08	
5	4.71	0.11	1.40	0.34	0.85	1.39	0.28	0.71	1.38	30.23	
5-1				0.27	0.49	1.89	0.27	0.45	2.06	15.66	
B. Investment-Grade Firms											
All				0.15	0.68	0.79	0.11	0.57	0.65	29.74	53.10
1	2.31	0.39	-1.42	0.07	0.60	0.38	0.02	0.50	0.14	30.56	
2	2.68	0.36	-0.45	0.09	0.62	0.51	0.05	0.52	0.32	29.66	
3	2.90	0.33	0.03	0.16	0.72	0.76	0.09	0.56	0.57	39.51	
4	3.09	0.11	0.48	0.18	0.75	0.83	0.14	0.65	0.73	24.89	
5	3.52	0.06	1.38	0.28	0.78	1.23	0.24	0.70	1.20	19.46	
5-1				0.21	0.38	1.94	0.22	0.37	2.13	5.19	
C. High-Yield Firms											
All				0.37	1.17	1.11	0.22	0.81	0.94	52.07	49.96
1	4.66	2.27	-1.34	0.12	0.87	0.49	0.03	0.64	0.14	45.88	
2	5.24	2.24	-0.48	0.24	1.14	0.73	0.11	0.82	0.44	48.26	
3	6.07	2.91	-0.03	0.53	1.52	1.20	0.30	1.00	1.06	56.72	
4	6.17	0.92	0.45	0.41	1.19	1.21	0.26	0.84	1.08	50.17	
5	6.92	0.37	1.45	0.58	1.47	1.36	0.41	1.00	1.42	53.72	
5-1				0.46	0.89	1.78	0.38	0.64	2.10	48.29	

This table contains five columns: (1) average values for actual spreads, model spreads, and pricing error (rank difference), (2) bond portfolio performance based on this error, (3) delta-hedging bond portfolio performance, (4) variance reduction (1 minus the variance ratio of hedged to unhedged investment excess returns), expressed as a percentage, and (5) average rank correlation between actual and model spreads for firms in the sample. For each portfolio, the table reports the average monthly excess returns over next month (R_p) in percentage, standard deviation (SD) in percentage, and annualized information ratio (IR).

Table 4: **Bias Corrections and Scale Adjustments for Time-Series and Cross-Sectional Analysis Using the Structural Model**

	Average				Unhedged			Hedged			Variance Reduction	Rank Correlation
Quintile	Actual	Pure Model	Debiasing Model	Pricing Error	Rp	SD	IR	Rp	SD	IR	(%)	(%)
A. All Firms												
All					0.21	0.76	0.97	0.14	0.59	0.79	39.73	89.34
1	3.07	0.45	4.04	-1.17	0.05	0.65	0.27	0.01	0.55	0.07	28.40	
2	3.57	0.62	3.93	-0.27	0.14	0.69	0.72	0.08	0.54	0.48	38.75	
3	4.53	1.10	4.65	-0.01	0.28	0.95	1.00	0.16	0.68	0.84	48.76	
4	4.19	0.73	3.95	0.26	0.27	0.86	1.08	0.18	0.67	0.95	39.31	
5	3.90	0.67	2.93	1.23	0.33	0.78	1.46	0.25	0.62	1.38	36.82	
5-1					0.28	0.38	2.52	0.24	0.28	2.95	45.71	
B. Investment-Grade Firms												
All					0.15	0.68	0.76	0.10	0.57	0.62	29.74	80.48
1	2.50	0.22	3.38	-1.23	0.02	0.61	0.11	-0.01	0.53	-0.04	24.51	
2	2.68	0.22	2.96	-0.31	0.09	0.65	0.48	0.06	0.55	0.36	28.40	
3	3.02	0.29	3.08	-0.01	0.17	0.76	0.78	0.11	0.61	0.62	35.58	
4	3.09	0.25	2.87	0.29	0.21	0.76	0.95	0.16	0.65	0.85	26.85	
5	3.12	0.25	2.28	1.30	0.25	0.69	1.27	0.20	0.59	1.16	26.89	
5-1					0.23	0.35	2.32	0.20	0.27	2.62	40.49	
C. High-Yield Firms												
All					0.35	1.13	1.06	0.20	0.79	0.89	51.12	84.50
1	5.02	1.35	6.20	-1.24	0.13	0.91	0.51	0.04	0.69	0.21	42.51	
2	5.36	1.46	5.89	-0.30	0.27	1.09	0.87	0.13	0.77	0.60	50.10	
3	6.41	2.05	6.59	0.01	0.44	1.44	1.05	0.25	0.96	0.90	55.56	
4	5.94	1.60	5.64	0.32	0.40	1.25	1.12	0.26	0.85	1.05	53.76	
5	5.94	1.71	4.89	1.31	0.51	1.25	1.41	0.35	0.90	1.35	48.16	
5-1					0.37	0.64	2.04	0.31	0.48	2.24	43.75	

This table contains five columns: (1) average values in percentage for actual spreads, model spreads, time-series debiased model spreads, and pricing error (rank difference between actual and time-series debiased model spreads), (2) bond portfolio performance based on this error, (3) delta-hedging bond portfolio performance, (4) variance reduction (1 minus the variance ratio of hedged to unhedged investment excess returns), expressed as a percentage, and (5) average rank correlation between actual and time-series debiased model spreads for firms in the sample. For each portfolio, the table reports the average monthly excess returns over next month (R_p) in percentage, standard deviation (SD) in percentage, and annualized information ratio (IR).

Table 5: **High- vs. Low-Liquidity Bond Portfolios**

	Average				Unhedged			Hedged			Variance Reduction	Rank Correlation
Quintile	Actual	Pure Model	Debiasing Model	Pricing Error	Rp	SD	IR	Rp	SD	IR	(%)	(%)
A. Firms with Higher Liquidity												
All					0.19	0.80	0.82	0.13	0.65	0.68	33.98	87.13
1	2.57	0.22	3.34	-1.20	0.09	0.72	0.44	0.06	0.60	0.33	30.56	
2	2.99	0.29	3.27	-0.27	0.15	0.79	0.66	0.10	0.63	0.56	36.40	
3	3.86	0.47	3.96	0.00	0.27	1.00	0.94	0.16	0.73	0.77	46.71	
4	3.43	0.32	3.22	0.28	0.22	0.90	0.84	0.15	0.72	0.72	36.00	
5	3.18	0.27	2.46	1.27	0.23	0.82	0.98	0.17	0.67	0.86	33.24	
5-1					0.14	0.40	1.20	0.11	0.27	1.38	54.44	
B. Firms with Lower Liquidity												
All					0.23	0.76	1.06	0.14	0.57	0.87	43.75	89.13
1	3.75	0.77	4.88	-1.17	-0.00	0.59	-0.02	-0.05	0.50	-0.32	28.18	
2	4.11	0.99	4.52	-0.29	0.13	0.73	0.64	0.06	0.56	0.35	41.15	
3	5.02	1.57	5.14	-0.00	0.32	1.03	1.06	0.18	0.70	0.90	53.81	
4	4.77	1.12	4.49	0.29	0.31	0.87	1.23	0.21	0.64	1.11	45.88	
5	4.72	1.12	3.54	1.25	0.43	0.81	1.84	0.33	0.63	1.84	39.51	
5-1					0.43	0.45	3.36	0.38	0.36	3.64	36.00	

This table contains five columns: (1) average values in percentage for actual spreads, model spreads, time-series debiased model spreads, and pricing error (rank difference between actual and time-series debiased model spreads), (2) bond portfolio performance based on this error, (3) delta-hedging bond portfolio performance, (4) variance reduction (1 minus the variance ratio of hedged to unhedged investment excess returns), expressed as a percentage, and (5) average rank correlation between actual and time-series debiased model spreads for firms in the sample. Panel A shows the bond manager trading bonds with higher liquidity, identified as those with above-median amounts of bonds outstanding. Panel B shows the bond manager trading bonds with lower liquidity, identified as those with below-median amounts of bonds outstanding. For each portfolio, the table reports the average monthly excess returns over next month (R_p) in percentage, standard deviation (SD) in percentage, and annualized information ratio (IR).

Table 6: **Evaluating Long-Short Bond Portfolio Performance over Different Holding Periods**

Holding Periods	Unhedged			Hedged			Variance Reduction
k	Rp	SD	IR	Rp	SD	IR	(%)
2	0.18	0.28	2.17	0.14	0.19	2.56	53.95
3	0.13	0.24	1.89	0.10	0.17	2.14	49.83
4	0.12	0.21	1.94	0.09	0.14	2.11	55.56
5	0.10	0.19	1.92	0.08	0.12	2.13	60.11
6	0.09	0.18	1.80	0.07	0.12	1.98	55.56
7	0.08	0.16	1.81	0.06	0.11	1.89	52.73
8	0.08	0.15	1.70	0.05	0.11	1.70	46.22

This table presents the performance of bond portfolios with holding periods ranging from 2 to 8 months. Portfolios are constructed monthly based on the chosen trading signal from pricing error and then held for k consecutive months, where $k = 2, 3, \ldots, 8$. For each portfolio, it reports the average monthly excess return over the next k months R_p in percentage, the standard deviation (SD) in percentage, and the annualized information ratio (IR).

Table 7: **Valuation and Hedging with Twisted Credit Spreads and Hedge Ratios**

		Cross-Sectional Debiasing									Combined Cross-Sectional and Time-Series Debiasing								
R		All	1	2	3	4	5	5-1	Variance Reduction 01	Variance Reduction 02	All	1	2	3	4	5	5-1	Variance Reduction 01	Variance Reduction 02
										A. All Firms									
0.5	Rp	0.06	-0.09	0.02	0.11	0.08	0.18	0.26			0.06	-0.03	0.01	0.06	0.12	0.14	0.17		
	SD	0.45	0.41	0.47	0.55	0.55	0.56	0.52	-33.53	-17.36	0.45	0.43	0.43	0.57	0.53	0.48	0.27	7.02	49.52
	IR	0.46	-0.73	0.16	0.67	0.52	1.09	1.76			0.46	-0.22	0.06	0.39	0.77	1.01	2.18		
0.6	Rp	0.09	-0.05	0.04	0.14	0.12	0.21	0.26			0.09	-0.01	0.03	0.11	0.15	0.18	0.19		
	SD	0.49	0.39	0.51	0.60	0.58	0.60	0.48	-13.78	0.00	0.49	0.46	0.47	0.62	0.57	0.50	0.24	26.53	60.11
	IR	0.63	-0.48	0.30	0.79	0.71	1.20	1.90			0.63	-0.09	0.25	0.61	0.89	1.23	2.71		
0.7	Rp	0.12	-0.02	0.07	0.17	0.15	0.24	0.26			0.12	0.00	0.06	0.15	0.18	0.22	0.21		
	SD	0.55	0.40	0.55	0.67	0.62	0.65	0.45	0.00	12.11	0.55	0.49	0.51	0.69	0.63	0.55	0.24	26.53	60.11
	IR	0.76	-0.19	0.42	0.87	0.86	1.27	1.99			0.76	0.02	0.42	0.76	0.97	1.37	3.01		
0.8	Rp	0.15	0.01	0.09	0.20	0.19	0.27	0.26			0.15	0.02	0.09	0.20	0.20	0.25	0.24		
	SD	0.61	0.46	0.61	0.74	0.68	0.71	0.45	0.00	12.11	0.61	0.54	0.58	0.76	0.70	0.61	0.27	7.02	49.52
	IR	0.85	0.07	0.52	0.93	0.97	1.31	2.02			0.85	0.12	0.53	0.89	1.02	1.44	2.98		
										B. Investment-Grade Firms									
0.5	Rp	0.03	-0.09	-0.02	0.02	0.08	0.17	0.27			0.03	-0.04	0.00	0.02	0.08	0.10	0.14		
	SD	0.43	0.41	0.41	0.45	0.53	0.60	0.53	-105.19	-94.53	0.43	0.43	0.41	0.45	0.56	0.47	0.29	-15.36	22.77
	IR	0.25	-0.80	-0.14	0.12	0.52	1.00	1.75			0.25	-0.34	0.03	0.13	0.50	0.74	1.68		
0.6	Rp	0.05	-0.06	0.00	0.04	0.10	0.19	0.25			0.05	-0.03	0.02	0.05	0.11	0.13	0.16		
	SD	0.46	0.39	0.43	0.46	0.56	0.63	0.47	-61.36	-52.98	0.46	0.45	0.44	0.49	0.57	0.48	0.24	20.99	47.11
	IR	0.41	-0.56	0.03	0.32	0.60	1.06	1.88			0.41	-0.23	0.16	0.36	0.65	0.93	2.27		
0.7	Rp	0.08	-0.03	0.02	0.07	0.12	0.21	0.24			0.08	-0.02	0.04	0.08	0.13	0.16	0.17		
	SD	0.50	0.40	0.45	0.50	0.61	0.66	0.42	-28.85	-22.16	0.50	0.47	0.48	0.54	0.60	0.51	0.21	39.51	59.50
	IR	0.54	-0.26	0.18	0.48	0.66	1.11	2.00			0.54	-0.13	0.28	0.54	0.76	1.07	2.82		
0.8	Rp	0.10	0.00	0.04	0.10	0.13	0.23	0.23			0.10	-0.01	0.06	0.12	0.16	0.19	0.19		
	SD	0.55	0.45	0.50	0.56	0.65	0.70	0.38	-5.48	0.0	0.55	0.51	0.53	0.60	0.65	0.55	0.22	33.61	55.56
	IR	0.63	0.02	0.31	0.60	0.71	1.15	2.06			0.63	-0.04	0.36	0.67	0.85	1.16	2.95		
										C. High-Yield Firms									
0.5	Rp	0.12	0.01	0.06	0.21	0.16	0.17	0.15			0.12	-0.02	0.05	0.14	0.21	0.24	0.25		
	SD	0.68	0.63	0.75	0.91	0.71	0.77	0.57	20.68	51.68	0.68	0.64	0.67	0.88	0.74	0.76	0.42	23.44	49.32
	IR	0.62	0.08	0.29	0.79	0.77	0.76	0.93			0.62	-0.09	0.26	0.56	0.98	1.08	2.10		
0.6	Rp	0.17	0.03	0.09	0.26	0.21	0.24	0.21			0.17	0.01	0.09	0.20	0.25	0.29	0.28		
	SD	0.75	0.67	0.80	0.98	0.76	0.80	0.50	38.96	62.82	0.75	0.68	0.73	0.94	0.81	0.82	0.41	27.04	51.71
	IR	0.77	0.18	0.41	0.91	0.94	1.06	1.46			0.77	0.06	0.43	0.76	1.08	1.23	2.37		
0.7	Rp	0.21	0.05	0.13	0.31	0.25	0.32	0.27			0.21	0.04	0.13	0.27	0.29	0.35	0.31		
	SD	0.83	0.71	0.87	1.07	0.85	0.89	0.49	41.38	64.29	0.83	0.73	0.80	1.03	0.90	0.91	0.43	19.75	46.88
	IR	0.89	0.27	0.50	1.00	1.04	1.25	1.87			0.89	0.20	0.56	0.89	1.13	1.34	2.50		
0.8	Rp	0.26	0.07	0.16	0.36	0.30	0.40	0.32			0.26	0.07	0.17	0.33	0.33	0.41	0.34		
	SD	0.92	0.76	0.95	1.17	0.95	1.03	0.56	23.44	53.36	0.92	0.79	0.88	1.16	1.00	1.00	0.47	4.12	36.54
	IR	0.97	0.34	0.58	1.07	1.10	1.33	2.00			0.97	0.31	0.66	0.98	1.15	1.40	2.49		

This table presents the results of cross-sectional debiasing and combined cross-sectional and time-series debiasing for bond portfolios with twisted credit spreads and hedge ratios. The analysis uses the recovery rate hyperparameter R, set to 0.5, 0.6, 0.7, and 0.8. Cross-sectional debiasing corrects for biases within each time period, while the combined approach further adjusts for biases across time periods. Variance reduction 01 is defined as 1 minus the ratio of the variance of hedged (twisted) to hedged (literal setting) investment excess returns, expressed as a percentage. Variance reduction 02 is defined as 1 minus the ratio of the variance of hedged (twisted) to unhedged (twisted) investment excess returns, expressed as a percentage. For each portfolio, the table reports the average monthly excess returns over next month (R_p) in percentage, the standard deviation (SD) in percentage, and the annualized information ratio (IR).

Table 8: **Robustness Checks with Leland-Toft Models with Varying Default Barriers.**

		Cross-Sectional Debiasing								Combined Cross-Sectional and Time-Series Debiasing							
d		All	1	2	3	4	5	5-1	Variance Reduction	All	1	2	3	4	5	5-1	Variance Reduction
A. Literal Setting (Leland-Toft Model)																	
0.01	Rp	0.23	0.07	0.13	0.26	0.30	0.38	0.31		0.22	0.06	0.15	0.29	0.27	0.32	0.26	
	SD	0.77	0.56	0.66	0.87	0.91	1.02	0.69		0.76	0.63	0.69	0.92	0.88	0.82	0.41	
	IR	1.03	0.44	0.69	1.02	1.16	1.29	1.55		0.98	0.33	0.74	1.08	1.07	1.35	2.18	
0.1	Rp	0.23	0.07	0.13	0.25	0.31	0.38	0.31		0.21	0.06	0.15	0.28	0.28	0.31	0.25	
	SD	0.76	0.56	0.66	0.87	0.91	1.02	0.68		0.76	0.63	0.69	0.91	0.88	0.81	0.41	
	IR	1.03	0.44	0.70	1.00	1.18	1.28	1.55		0.98	0.33	0.73	1.06	1.10	1.33	2.10	
0.3	Rp	0.23	0.07	0.13	0.25	0.30	0.38	0.31		0.21	0.05	0.14	0.28	0.27	0.30	0.25	
	SD	0.76	0.56	0.65	0.88	0.89	1.03	0.70		0.76	0.63	0.69	0.91	0.88	0.81	0.41	
	IR	1.03	0.46	0.70	0.97	1.17	1.28	1.52		0.98	0.33	0.73	1.06	1.10	1.33	2.10	
B. Literal Setting with Delta-Hedging																	
0.01	Rp	0.09	-0.01	-0.01	0.01	0.20	0.29	0.30		0.11	0.01	0.06	0.13	0.15	0.19	0.17	
	SD	0.61	0.60	0.71	0.98	0.69	0.79	0.68	2.59	0.60	0.55	0.57	0.72	0.67	0.71	0.39	8.01
	IR	0.53	-0.07	-0.04	0.03	0.99	1.26	1.52		0.62	0.09	0.39	0.64	0.75	0.91	1.52	
0.1	Rp	-0.12	-0.14	-0.35	-0.47	0.05	0.29	0.43		-0.03	-0.01	-0.08	-0.07	-0.01	0.03	0.04	
	SD	1.39	1.63	2.57	2.54	1.60	0.78	1.76	-558.16	0.93	0.71	1.17	1.15	1.07	1.09	0.72	-200.14
	IR	-0.31	-0.30	-0.48	-0.64	0.12	1.30	0.85		-0.10	-0.05	-0.24	-0.21	-0.04	0.11	0.21	
0.3	Rp	-0.94	-0.48	-1.41	-2.13	-0.58	-0.12	0.37		-0.37	-0.18	-0.28	-0.58	-0.51	-0.30	-0.12	
	SD	4.16	3.80	6.50	7.86	3.19	2.15	3.50	-2415.14	1.78	1.39	1.80	2.55	2.33	1.81	1.20	-747.74
	IR	-0.78	-0.44	-0.75	-0.94	-0.63	-0.19	0.36		-0.72	-0.45	-0.54	-0.78	-0.75	-0.58	-0.35	
C. CDS-Based Setting (Leland-Toft Model)																	
0.7	Rp	0.18	0.06	0.11	0.25	0.21	0.29	0.23		0.17	0.02	0.09	0.25	0.22	0.28	0.26	
	SD	0.80	0.71	0.76	0.95	0.81	0.91	0.48		0.80	0.68	0.73	0.99	0.87	0.84	0.40	
	IR	0.79	0.30	0.51	0.90	0.90	1.09	1.64		0.75	0.11	0.44	0.88	0.88	1.16	2.24	
0.8	Rp	0.18	0.06	0.11	0.26	0.21	0.29	0.23		0.17	0.02	0.10	0.25	0.23	0.28	0.26	
	SD	0.80	0.70	0.79	0.93	0.84	0.88	0.45		0.79	0.68	0.74	0.97	0.89	0.83	0.38	
	IR	0.80	0.28	0.47	0.97	0.86	1.13	1.76		0.75	0.08	0.45	0.89	0.88	1.16	2.39	
0.9	Rp	0.18	0.06	0.10	0.27	0.21	0.27	0.22		0.17	0.02	0.10	0.25	0.23	0.28	0.26	
	SD	0.79	0.74	0.76	0.95	0.86	0.82	0.40		0.79	0.66	0.75	0.96	0.89	0.83	0.38	
	IR	0.80	0.27	0.47	1.00	0.84	1.16	1.87		0.76	0.09	0.46	0.89	0.89	1.15	2.38	
D. CDS-Based Setting with Delta-Hedging																	
0.7	Rp	0.09	-0.05	0.04	0.14	0.13	0.21	0.26		0.10	-0.03	0.04	0.14	0.15	0.19	0.22	
	SD	0.65	0.59	0.64	0.76	0.68	0.77	0.50	-7.37	0.67	0.60	0.63	0.80	0.73	0.69	0.36	21.06
	IR	0.49	-0.32	0.20	0.64	0.68	0.94	1.82		0.51	-0.15	0.21	0.62	0.72	0.95	2.10	
0.8	Rp	0.08	-0.05	0.03	0.14	0.12	0.19	0.25		0.08	-0.05	0.01	0.12	0.13	0.18	0.23	
	SD	0.64	0.54	0.66	0.72	0.69	0.71	0.43	9.30	0.64	0.57	0.60	0.76	0.70	0.66	0.31	34.20
	IR	0.46	-0.33	0.13	0.66	0.61	0.94	1.99		0.42	-0.32	0.08	0.56	0.63	0.92	2.56	
0.9	Rp	0.07	-0.06	0.01	0.09	0.11	0.18	0.24		0.05	-0.06	0.01	0.03	0.12	0.16	0.21	
	SD	0.72	0.55	0.62	1.58	0.69	0.66	0.47	-39.13	0.72	0.53	0.60	1.54	0.67	0.63	0.31	32.60
	IR	0.32	-0.36	0.07	0.21	0.54	0.94	1.74		0.24	-0.36	0.04	0.07	0.60	0.87	2.38	

This table presents the investment performance based on the Leland-Toft model, comparing two distinct settings: the literal setting and the CDS-based setting. Each setting is evaluated under two debiasing approaches—cross-sectional debiasing and combined cross-sectional and time-series debiasing —across different default barrier levels. Panels A and B focus on the literal setting, where the default barriers are set at $d = 0.01, 0.1,$ and 0.3. Panel A reports the investment performance, which reflects the model's ability to capture excess returns under the specified default barriers and debiasing methods. Panel B evaluates the delta-hedging performance in the same setting. Panels C and D focus on the CDS-based setting, where the recovery rate is fixed at $R = 0.7$ and the default barriers are set at $d = 0.7, 0.8,$ and 0.9. Similar to the literal setting, Panel C presents the investment performance, while Panel D assesses the delta-hedging performance. Variance reduction is defined as 1 minus the ratio of the variance of hedged to unhedged investment excess returns, expressed as a percentage. For each portfolio, the table reports three key performance metrics: the average monthly excess returns over next month (R_p) in percentage, the standard deviation (SD) in percentage, and the annualized information ratio (IR).

7 Appendix

7.1 Implementation of Leland and Toft (1996)

Leland and Toft (1996) extends the Black and Cox (1976) framework by assuming that the firm continuously pays coupons at a constant rate $\bar{c}$, with each payment equal to $c = \bar{c}D$, where D denotes the bond's face value. Coupon payments stop immediately upon default, whether it occurs before or at maturity, defined as:

$$\tau^{\text{def}} = \inf\{s \in (t,T] : A_s \leq K_s\}, \quad K_s = dD \cdot \mathbf{1}_{\{t<s<T\}} + D \cdot \mathbf{1}_{\{s=T\}},$$

where K_s is the default barrier and $d \in (0,1)$ is a constant scaling parameter. The model yields the following pair of conditions linking the firm's equity value and equity return volatility to its asset value and asset return volatility:

$$E = E^{\text{Leland-Toft}}, \tag{11}$$

$$\sigma_E = (A/E)\sigma_A \partial E^{\text{Leland-Toft}}/\partial A, \tag{12}$$

where

$$E^{\text{Leland-Toft}} = A - B^{\text{Leland-Toft}} = A - \left(B^{\text{Black-Cox}} + B^{\text{Coupon}}\right),$$

and

$$B^{\text{Black-Cox}} = A - E^{\text{Black-Cox}} = A - \Big(BS(A,D,\tau,r,\sigma) - AH^{2r/\sigma^2} BS(H,1/d,\tau,r,\sigma)\Big),$$

$$B^{\text{Coupon}} = \mathbb{E}^{\mathbb{Q}}\Big(\int_t^T ce^{-(s-t)}\mathbf{1}_{\{\tau^{\text{def}}\geq s\}} ds \Big| \mathbb{F}_t\Big),$$

$$= (c/r)\left[1 - e^{-r\tau}\left(N(h_1(A,\tau)) - H^{2\widetilde{a}} N(h_2(A,\tau))\right) - H^{\widetilde{a}+\xi} N(h_3(A,\tau)) - H^{\widetilde{a}-\xi} N(h_4(A,\tau))\right].$$

The derivative of the equity value with respect to the asset value is given by

$$\partial E^{\text{Leland-Toft}}/\partial A$$

$$= \partial E^{\text{Black-Cox}}/\partial A - \partial B^{\text{Coupon}}/\partial A$$

$$= N(\delta_1(A/D,\tau))$$

$$+ H^{2r/\sigma^2+1} N(\delta_1(dH,\tau)) - \left(1 - 2r/\sigma^2\right) H^{2r/\sigma^2} BS(H,1/d,\tau,r,\sigma)$$

$$- (c/r) A^{-1}\left\{-e^{-r\tau}\left(\left(\phi(h_1(A,\tau)) + H^{2\widetilde{a}}\phi(h_2(A,\tau))\right) \Big/ \left(\sigma\sqrt{\tau}\right) + H^{2\widetilde{a}}(2\widetilde{a}) N(h_2(A,\tau))\right)\right.$$

$$\left. + \left(H^{\widetilde{a}+\xi}(\widetilde{a}+\xi) N(h_3(A,\tau)) + H^{\widetilde{a}-\xi}(\widetilde{a}-\xi) N(h_4(A,\tau))\right)\right.$$

$$+ \left(H^{\widetilde{a}+\xi}\phi\left(h_3\left(A,\tau\right)\right) + H^{\widetilde{a}-\xi}\phi\left(h_4\left(A,\tau\right)\right)\right) \Big/ \left(\sigma\sqrt{\tau}\right)\Big\},$$

where

$$\tau = T - t,\ H = dD/A,\ \theta = \widetilde{a} + 1,$$

$$\widetilde{a} = \sigma^{-2}\left(r - 0.5\sigma^2\right),\ \xi = \sigma^{-2}\sqrt{\left(r - 0.5\sigma^2\right)^2 + 2r\sigma^2},$$

$$BS\left(A, D, \tau, r, \sigma\right) = AN\left(\delta_1\left(A/D, \tau\right)\right) - De^{-r\tau}N\left(\delta_2\left(A/D, \tau\right)\right),$$

$$\delta_1\left(s, \tau\right) = \left(\ln(s) + \left(r + 0.5\sigma^2\right)\tau\right) / \left(\sigma\sqrt{\tau}\right),$$

$$\delta_2\left(s, \tau\right) = \left(\ln(s) + \left(r - 0.5\sigma^2\right)\tau\right) / \left(\sigma\sqrt{\tau}\right),$$

$$h_1\left(A, \tau\right) = \left(\ln\left(A/\left(dD\right)\right) + \left(r - 0.5\sigma^2\right)\tau\right) / \left(\sigma\sqrt{\tau}\right),$$

$$h_2\left(A, \tau\right) = \left(\ln\left(dD/A\right) + \left(r - 0.5\sigma^2\right)\tau\right) / \left(\sigma\sqrt{\tau}\right),$$

$$h_3(A, \tau) = \left(\ln\left(dD/A\right) + \sigma^2\,\xi\,\tau\right) / \left(\sigma\sqrt{\tau}\right),$$

$$h_4(A, \tau) = \left(\ln\left(dD/A\right) - \sigma^2\,\xi\,\tau\right) / \left(\sigma\sqrt{\tau}\right).$$

For **Equation** (11) and (12), the involved variables are: coupon rate (c), default barrier (d), asset value (A), asset return volatility (σ_A), equity value (E), equity return volatility (σ_E), principal payment (D), debt maturity (T), and interest rate (r). The variables asset value (A) and asset return volatility (σ_A) require estimation by using **Equation** (11) and (12). The remaining variables can be directly tied to available data:

- Coupon rate $(\bar{c})$: I calculate the annualized coupon rate for each bond using data from the WRDS Bond Database and aggregate it to the firm level using bond outstanding weights. Because Leland and Toft (1996) assumes continuous coupon payments while companies typically pay semi-annually. To address this, I divide the annualized coupon rate by 365.

- Default barrier (d): This is a default barrier from the Black-Cox model defined by a proportional constant, which triggers firm default when the asset value falls below $d \times D$. I fine-tuned this hyperparameter by testing various values within a predefined range to evaluate the model's performance.

- The remaining inputs follow Merton (1974), as discussed in Section 4.

The default probability formula in Leland and Toft (1996) is mathematically identical to that in Black and Cox (1976). However, because Leland and Toft (1996) incorporates continuous coupon payments, the estimated asset value, its volatility, and consequently the model-implied default probabilities may differ. This framework also contrasts with Geske (1977), which assumes that default can occur at coupon or maturity payment dates if the firm's asset value falls short of the required payment. Under the risk-neutral measure $\mathbb{Q}$, the default probability in Leland and Toft (1996) is given by

$$\begin{aligned}
&\mathbb{Q}^{\text{Leland-Toft}}(A, \sigma_A, D, t, T, r, d) \\
&= \mathbb{Q}(\{\min_{t<s<T} A_s \leq dD\}) + \mathbb{Q}(\{\min_{t<s<T} A_s > dD,\ A_T < D\}) \\
&= N\big((\ln(D/A) - (r - 0.5\sigma^2)\tau)/(\sigma\sqrt{\tau})\big) \\
&\quad + (A/(dD))^{1-2r/\sigma^2}\, N\big((\ln(d^2 D/A) + (r - 0.5\sigma^2)\tau)/(\sigma\sqrt{\tau})\big),
\end{aligned}$$

where I consider two components of the default probability: (i) the firm defaults if its asset value falls below the default boundary before maturity, and (ii) the firm defaults if its asset value falls below total debt at maturity. I use this decomposition to compare performance with the default probability implied by the Merton model.

The credit spread implied by the Leland-Toft model can be expressed as:

$$cs^{\text{LTCSB}}(A, \sigma_A, D, t, T, r, d, c) = -\ln\left(\left(De^{-r\tau}\right)^{-1} B^{\text{Leland-Toft}}\right) \Big/ \tau, \tag{13}$$

$$cs^{\text{LTCSC}}(A, \sigma_A, D, t, T, r, d, c, R) = -\ln\left(1 - \mathbb{Q}^{\text{Leland-Toft}}\right)(1 - R) \Big/ \tau. \tag{14}$$

Equation (13) presents a literal formulation of the **L**eland–**T**oft **C**redit **S**pread (LTCSB), derived from the market **B**ond value. Equation (14) provides an alternative representation based on the survival probability—defined as one minus the Leland–Toft default probability—under the constant hazard rate assumption within the fair **C**DS spread framework, incorporating a recovery rate R.

For the hedge ratio implied by the Leland-Toft model, I use $\partial B/\partial A$. The following

hedge ratios with respect to asset value:

$$h_B^{\text{LTHRB}}(A, \sigma_A, B, D, t, T, r, d, c) = 1 - \partial E^{\text{Leland-Toft}}/\partial A, \tag{15}$$

$$h_B^{\text{LTHRC}}(A, \sigma_A, B, D, t, T, r, d, c, R) = -B(1-R)\left(\partial \mathbb{Q}^{\text{Leland-Toft}}/\partial A\right) \Big/ \left(1 - \mathbb{Q}^{\text{Leland-Toft}}\right), \tag{16}$$

where

$$\begin{aligned}
&\partial \mathbb{Q}^{\text{Leland-Toft}}/\partial A \\
&= -\phi\Big(\left(\ln\left(D/A\right) - \left(r - 0.5\sigma^2\right)\tau\right)/\left(\sigma\sqrt{\tau}\right)\Big) \Big/ \left(\sigma\sqrt{\tau}A\right) \\
&\quad + \left(1 - 2r/\sigma^2\right)\left(A/\left(dD\right)\right)^{-2r/\sigma^2} N\Big(\left(\ln\left(d^2D/A\right) + \left(r - 0.5\,\sigma^2\right)\tau\right)/\left(\sigma\sqrt{\tau}\right)\Big) \Big/ \left(dD\right) \\
&\quad - \left(A/\left(dD\right)\right)^{1-2r/\sigma^2} \phi\Big(\left(\ln(d^2D/A) + (r - 0.5\,\sigma^2)\tau\right)/(\sigma\sqrt{\tau})\Big) \Big/ \left(\sigma\sqrt{\tau}\right).
\end{aligned}$$

Equation (15) represents the literal formulation of the **L**eland–**T**oft **H**edge **R**atio with respect to the asset value, derived by differentiating the market value of the **B**ond, termed **LTHRB**. In contrast, Equation (16) provides an alternative hedge ratio by decomposing $\partial B/\partial A$ as $(\partial B/\partial CS)(\partial CS/\partial A) = (\partial B/\partial CS_1)(\partial CS_2/\partial A)$, where CS_1 is the literal credit spread given in Equation (13) and CS_2 is the credit spread under the **C**DS framework from Equation (14), termed **LTHRC**. These Leland-Toft model settings are utilized to facilitate a comparison with the Merton model in Section 5.4, with respect to valuation and hedging.