FINM 32900 Final Project: Replicating Tables from *Noisy Prices and Return-based Anomalies in Corporate Bonds*

Yuxuan (Ryan) Bai, Ximeng Deng, Yitong Wang,
Yijing Zhang, Siying Zheng

The University of Chicago

February 2024

Abstract

This study endeavors to replicate Table 1 from the research article "Noisy Prices and Return-based Anomalies in Corporate Bonds" authored by Alexander Dickerson, Cesare Robotti, and Giulio Rossetti. Utilizing the TRACE database, we extracted data on corporate bonds and computed various metrics, including bid-ask bias, return, bid-ask spread, and credit spread, across different time periods and bond ratings. Our findings indicate that the replication results for Table 1 closely align with the statistics reported in the original study. However, it is imperative to acknowledge several challenges encountered during our analysis, which may have contributed to slight discrepancies between our results and those presented in the original paper.

1 Introduction

1.1 Background

The purpose of this project is to replicate table 1 from the paper "Noisy Prices and Return-based Anomalies in Corporate Bonds" written by Alexander Dickerson, Cesare Robotti, and Giulio Rossetti. As demonstrated by Dickerson et al. (2023),this paper tries to question the prevailing view that investors can earn usually high returns from certain patterns in corporate bond prices, such as return reversals and momentum. The authors debate these large gains are most likely a result of two factors: market microstructure noise an the use of arbitrary methods to limit extreme values in return data.

1.2 Objective

Our object is to replicate table 1 from page 11 from the paper. Table 1 in the paper is critical as it sets the foundation for the study's argument by presenting an analysis of bid-ask bias over time in the corporate bond market. The table includes time-series averages of cross-sectional means for daily bid-ask bias, return, bid-ask spread, and credit spread,

all measured in basis points (bps). This information is crucial for the paper's thesis because it illustrates the extent of market microstructure noise — discrepancies in bond prices due to the mechanics of trading rather than changes in fundamentals — and how it can influence returns.

2 Methodology

2.1 Data Sources

The data source of Table 1 is the Wharton Research Data Services (WRDS) bond database, which integrates information from the Enhanced Trade Reporting and Compliance Engine (TRACE). We also utilize data from Mergent Fixed Income Securities Database (FISD) for details on bond issues.

We followed the guidelines from OpenBondAssetPricing (2023) website, specifically using the code from Dickerson (2023) to generate our final replicated table. We used "MakeBondIntra_Daily.py" (renamed as "load_trace.py") to generate the price and volume data.

We apply the following filters to select bonds from WRDS bond database:

1. Remove investment grade (IG) rated bonds that have less than USD 150 million outstanding prior to, and including, November 2004, and less than USD 250 million after November 2004.

2. Remove non-investment grade (HY) rated bonds that have less than USD 100 million outstanding prior to, and including, September 2016, and less than USD 250 million after September 2016.

3. Remove bonds that are classified as zero-coupon, bond type == 'CMTZ'.

4. Remove bonds that are classified as convertible, conv == 'N'.

We merge the WRDS data with FISD (also publicly available via the WRDS data platform) and apply the following filters that are all standard in the literature:

1. Only keep bonds that are issued by firms domiciled in the United States of America, COUNTRY DOMICILE == 'USA'.

2. Remove bonds that are private placements, PRIVATE PLACEMENT == 'N'.

3. Only keep bonds that are traded in U.S. Dollars, FOREIGN CURRENCY == 'N'.

4. Bonds that trade under the 144A Rule are discarded, RULE 144A == 'N'.

5. Remove all asset-backed bonds, ASSET BACKED == 'N'.

6. Remove convertible bonds, CONVERTIBLE == 'N'.

7. Only keep bonds with a fixed or zero coupon payment structure, i.e., remove bonds with a floating (variable) coupon, COUPON TYPE != 'V'.

8. Remove bonds that are equity linked, agency-backed, U.S. Government, and mortgage-backed, based on their BOND TYPE.

9. Remove bonds that have a "nonstandard" interest payment structure or bonds not caught by the variable coupon filter (COUPON TYPE). This affects a tiny fraction of bonds (0.10% or 142 bonds) of the FISD data file. We remove bonds that have an INTEREST FREQUENCY equal to -1 (N/A), 13 (Variable Coupon), 14 (Bi-Monthly), and 15 and 16 (undocumented by FISD). Additional information on INTEREST FREQUENCY is available on Page 60 of 67 of the FISD Data Dictionary 2012 document.

10. Remove a small fraction of bonds that do not have the required (and crucial information) to compute accrued interest. Bonds that do not have a valid DATED DATE are removed (3,051 bonds). The DATED DATE variable is the date from which bond interest accrues. Bonds without a valid INTEREST FREQUENCY, DAY COUNT BASIS, OFFERING DATE, COUPON TYPE, and COUPON are also removed (425 bonds in total)

To adjust the change in data structure in TRACE data, We apply the following filters in cleaning the intraday TRACE data for the pre-2012 database:

1. Keep all trades that have less than two days to settlement, days_to_sttl_ct == '002', days_to_sttl_ct == '001', days_to_sttl_ct == '000' or days_to_sttl_ct == 'None'.

2. Remove trade records with the 'when-issued' indicator, wis_fl != 'Y'.

3. Remove trade records with the 'locked-in' indicator, lckd_in_ind != 'Y'.

4. Keep trade records which do not have special conditions, sale_cndtn_cd == 'None' or sale_cndtn_cd == '@'.

When we generate the corporate daily returns, here is the criteria we followed to filter our data to align with the assumptions mentioned from the paper:

1. Exclude observations with more than 5 business days since the last trade. This ensures that the calculated daily returns are based on recent and valid data.

2. Filter out bonds with fewer than five trades per month. This filtration is achieved by tabulating the counts of each bond's trades within each month.

3. Compute the daily returns, remove significant return reversals, and exclude returns with an absolute value exceeding 20%.

2.2 Analysis Methods

2.2.1 Bid-ask Bias

For calculating bid-ask bias, we utilized price and volume data loaded from trace and assume that each trade occur with equal probability. For each day d and for each bond i, we first compute the volume-weighted average bid ($\overline{P}_B$) and ask ($\overline{P}_A$) prices with trades that have a volume greater than $10,000$. We then approximate the bid-ask bias $\sigma^2[\delta]$ using the following equation:

$$\sigma^2[\delta_i] = \left(\frac{\overline{P}_A - \overline{P}_B}{\overline{P}_A + \overline{P}_B}\right)^2 \tag{1}$$

To address extreme outliers, we winsorize the distribution of the bid-ask bias at the 99.5% level.

2.2.2 Return

For calculating returns, we utilized data generated from `MakeBondDailyMetrics.py` to derive the daily return of each bond, referred to as `BondDailyPublic`. This dataset encompasses various information such as issue date and credit spread. However, for our analysis, we require only the issue date, CUSIP ID, and clean price to compute the bond's daily return. The process of data cleaning is outlined in `Calculate_bond_return.ipynb`.

The formula we used to calculate the daily bond return is simply:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{2}$$

Where R_t is the return on the day, P_t is the price of the bond on the current day, P_{t-1} is the price of the bond on the prrevious day.

Following the application of these criteria, we have generated a dataset suitable for further analysis, including calculating the mean of daily returns based on different time horizons and credit ratings.

2.2.3 Bid-ask spread

For calculating bid-ask spread, we use the same volume-weighted price calculated in bid-ask bias. The bid-ask spread is defined as the difference between the daily volume-weighted average bid and ask prices scaled by their average.

$$bid - ask\ spread = 2 * \left(\frac{\overline{P}_B - \overline{P}_A}{\overline{P}_A + \overline{P}_B} \right), \tag{3}$$

2.2.4 Credit Spread

Similar with return, when calculating credit spread, we also utilized data generated from `MakeBondDailyMetrics.py` to derive the daily credit spread of each bond, referred to as `BondDailyPublic`. Specifically in these files, the credit spread is calculated by the difference between daily bond yields and a duration-matched portfolio of U.S. Treasury Bonds, and convert the result into basis points(bps).

Initial result shows that there has been extreme outliers in the sample. Therefore we winsorize the distribution of the credit spread at the 99.5% level.

2.2.5 Rating

In our approach, we initiated by establishing a connection to the WRDS database to procure ratings information, specifically honing in on the ratings provided by Moody's and Standard & Poor's (SP), while excluding other types. We assigned numerical values to each rating from SP and Moody's to standardize the data. These numerical values were further categorized into three broad classifications: 'A and above', 'BBB', or 'Junk', simplifying the analysis of credit quality.

Our script contains dedicated functions for processing SP and Moody's ratings individually, utilizing a mapping strategy to convert textual ratings into numeric values, which are then categorized based on predetermined thresholds. We took steps to ensure data integrity by filtering out duplicates and irrelevant ratings such as 'NR' (not rated), 'SUSP' (suspended), and others, focusing our attention on pertinent and accurate data.

Furthermore, we combined Moody's and SP ratings, ensuring that in instances where SP ratings were absent, they were replaced with Moody's ratings. This approach was critical for maintaining the completeness and accuracy of our dataset. Following this, we conducted a filtration process to remove certain non-relevant ratings, ensuring the dataset's relevancy and reliability.

To conclude, we organized the data by CUSIP (a unique identifier for financial instruments) and rating date, creating a structured and chronological dataset. This organized data was then saved to a CSV file, rendering it ready for further analysis or reporting. Our process embodies a thorough method for cleaning, categorizing, and preparing credit ratings data, prioritizing accuracy, relevance, and user-friendliness.

2.3 Summary Statistics

	spread	winsorized_bias	daily_return_bps	cs_dur_bps
count	5285.000000	5285.000000	5285.000000	5285.000000
mean	33.968771	0.138678	1.929342	329.097031
std	21.671959	0.150115	35.866739	204.175992
min	-13.377713	0.000000	-410.471986	-5.575916
25%	18.103552	0.034829	-9.374882	208.482887
50%	30.352161	0.086850	2.960880	285.749386
75%	43.666566	0.185730	13.776846	388.669565
max	272.164555	1.834768	1250.000068	6053.541533

Table 1: Summary Statistics

The provided table outlines a set of summary statistics for bond market data, encompassing 5285 observations across four key metrics: spread, winsorized bias, daily return in basis points (bps), and credit spread duration in bps (cs_dur_bps). The average spread is noted at approximately 33.97 bps, with a standard deviation indicating moderate variability. Winsorized bias is low on average, around 0.14, with minimal variance, suggesting that extreme values are limited and data concentration is near the mean. Daily returns average at about 1.92 bps but exhibit high variability, as evidenced by a substantial standard deviation of roughly 35.87. The cs_dur_bps averages at 329.09 with a similarly high standard deviation, pointing to a wide distribution around the mean.

Examining the distribution, the minimum spread recorded is negative, at -13.38 bps, while the maximum reaches 272.16 bps, indicating a broad spread range. The daily return has a notably wide range, plunging to a minimum of -410.48 bps and soaring to a maximum of 1250 bps, highlighting the potential for extreme negative and positive returns in the dataset. The percentile values for both spread and cs_dur_bps suggest a progressive widening from the lower to the upper quartiles. The maximum value for cs_dur_bps is particularly high at 6053.54, which, alongside the high maximum for daily returns, may signal the presence of outliers or a skewed distribution with a long tail to the right.

category	variables	All	A and above	BBB	Junk
Full sample	spread	28.767	25.584	24.529	36.097
	winsorized_bias	0.104	0.088	0.080	0.143
	daily_return_bps	1.093	0.151	0.574	2.460
	cs_dur_bps	299.549	114.333	169.053	599.449
Pre-crisis	spread	50.237	38.653	53.043	62.748
	winsorized_bias	0.253	0.173	0.272	0.339
	daily_return_bps	1.847	-0.413	-0.682	6.386
	cs_dur_bps	331.988	96.097	202.283	714.090
Crisis	spread	59.388	53.730	60.430	67.507
	winsorized_bias	0.304	0.242	0.325	0.387
	daily_return_bps	-1.731	-0.439	5.751	-9.251
	cs_dur_bps	660.251	286.105	558.881	1322.028
Post-Crisis	spread	39.384	32.847	39.809	44.822
	winsorized_bias	0.149	0.105	0.148	0.188
	daily_return_bps	6.944	3.742	5.467	11.007
	cs_dur_bps	442.492	161.060	249.777	851.743
Basel II.5 & III	spread	27.007	20.208	25.189	33.745
	winsorized_bias	0.068	0.044	0.058	0.095
	daily_return_bps	1.645	0.200	1.342	2.988
	cs_dur_bps	288.978	102.203	176.537	541.398
Post-Volcker	spread	18.235	12.791	16.220	24.695
	winsorized_bias	0.043	0.023	0.033	0.068
	daily_return_bps	-0.397	-0.765	-0.499	-0.005
	cs_dur_bps	222.882	75.045	130.228	447.545
Up to latest	spread	28.420	25.285	24.175	35.788
	winsorized_bias	0.102	0.087	0.078	0.141
	daily_return_bps	1.155	0.208	0.683	2.484
	cs_dur_bps	296.450	113.624	167.989	594.531

Table 2: Replication of Table 1 from the Paper

3 Results

3.1 Comparison with Original Work

Comparing our replicated table (see Table 2: Replication of Table 1 from the Paper) to the original study's result, our results are off but not in a significant level. We concluded the reasons that we cannot perfectly replicate the figures:

1. **Varying Approaches Yield Divergent Outcomes**. The paper lacks comprehensive details necessary for replicating the process, allowing for a degree of subjectivity in data loading and analytical methods.

2. **Utilization of an Outdated Dataset**. The dataset employed in our replication may be more current than the one used by the original authors. It is plausible that the WRDS and TRACE databases have since updated their historical records, potentially excluding anomalies or erroneous bond data. Consequently, the discrepancy in datasets could account for the variations in our results compared to the

original study.

3. **Exclusion of Maturity Information in the Filtering Criteria**. Our data filtering process omitted consideration of bond maturities. Despite implementing the three specified filtering conditions from section 2.1, we overlooked the criterion that mandates the exclusion of bonds with maturities under one year.

4 Discussion

4.1 Successes

Although the replication process was demanding, we achieved several significant milestones:

1. We successfully replicated Table 1, with figures that closely align with the original results.

2. We gained a comprehensive understanding of the methodologies used to calculate various bond metrics.

3. We thoroughly examined the authors' coding process by reviewing their GitHub repository.

4. We met all the requirements set forth by the FINM 32900 Final Project, ensuring that our project was thorough, comprehensive, and well-documented.

4.2 Challenges

Although our replication did demonstrated our solid understanding of the paper, we encounter many challenges and we managed to address most of them.

1. **Absence of Access to a Database System.** One prominent issue involved the handling of large datasets, where conventional methods such as basic Python or pandas read operations proved to be time-consuming.

2. **Hardware Inefficiency.** Personal laptops are used for the entire replication process, which is less efficient compared to industrial level hardware. Especially when constructing daily return and credit spread, the original code crashed the kernel on the laptop. Our replication process would be much smoother with more efficient processing tools accessible.

3. **Imprecise Analysis.** Given the limited time to finish this project, we only have about two weeks downloading the dataset and replicating the table. The original authors omitted certain steps that, although seemingly trivial, necessitate experimental adjustments and independent exploration on our part. The restricted time frame limits our ability to ensure the utmost accuracy of our results. If more time allowed, the result would be more precise.

4. **Time Cost for Each Step.** The efficiency of each step is constrained by limited data sources and hardware resources, impacting our ability to download and process data swiftly. The initial steps of downloading the data and generating return and

credit spread alone consume nearly an entire day. This significantly slows the progress of the project.

5 Conclusion

In conclusion, our project to replicate Table 1 from "Noisy Prices and Return-based Anomalies in Corporate Bonds" was met with both successes and challenges, which provided invaluable learning experiences. Our replication closely mirrored the original results, affirming the robustness of the methodologies employed by Dickerson, Robotti, and Rossetti. Despite encountering certain obstacles, such as data discrepancies due to updates and the absence of detailed replication guidelines, our endeavor demonstrates that with meticulous methodology and diligent effort, critical academic findings can be substantiated.

References

Dickerson, A. (2023). Trace corporate bond processing. `https://github.com/Alexander-M-Dickerson/TRACE-corporate-bond-processing`. Accessed: 2024-02-01.

Dickerson, A., Robotti, C., and Rossetti, G. (2023). Noisy prices and return-based anomalies in corporate bonds. *UNSW Business School Research Paper Forthcoming*. Available at SSRN: https://ssrn.com/abstract=4575879.

OpenBondAssetPricing (2023). Daily bond return and credit spread data. OpenBondAssetPricing website. Accessed: 2024-02-01.