Pricing of Corporate Bonds: Evidence From a Century-Long Cross-Section*

Mohammad Ghaderi[†] Sébastien Plante[‡] Nikolai Roussanov[§] Sang Byung Seo[¶]

February 23, 2025

Abstract

We construct a new historical corporate bond database spanning 128 calendar years to address longstanding data limitations in corporate bond research. By hand-collecting monthly corporate bond quotes from three archival print sources, we complement existing datasets and create an extensive database dating back to 1895, comprising nearly 110,000 unique bonds and 8 million observations. Leveraging this expanded sample, we find that the lack of priced risks in corporate bonds documented by recent studies stems from their reliance on short samples. With greater statistical power, we show that prominent bond and stock factors as well as nontraded macroeconomic factors are significantly priced with theoretically consistent signs. Our database, covering major economic episodes like the Great Depression, not only validates previous empirical findings like the GZ spread's predictive content but also aims to serve as a "CRSP" for corporate bonds.

*We thank Francesca Carapella, Yoshio Nozawa, and seminar participants at the Federal Reserve Board for helpful comments and suggestions. Support for this research was provided by the University of Wisconsin-Madison, Office of the Vice Chancellor for Research with funding from the Wisconsin Alumni Research Foundation. We also gratefully acknowledge the financial support from the Macro Finance Society.

[†]School of Business, University of Kansas; mghaderi@ku.edu

[‡]Wisconsin School of Business, University of Wisconsin-Madison; splante@wisc.edu

[§]The Wharton School, University of Pennsylvania and NBER; nroussan@wharton.upenn.edu

[¶]Wisconsin School of Business, University of Wisconsin-Madison; sang.seo@wisc.edu

1 Introduction

The corporate bond market, a cornerstone of capital markets, plays a critical role alongside the stock market in raising capital and financing long-term investments for firms. Despite its significance in the real world, the academic literature has predominantly focused on stocks and their returns over the past half-century. This imbalance is largely driven by data availability: unlike stocks, which benefit from the availability of a comprehensive database, no equivalent long-term dataset exists for corporate bonds. The complexity of corporate bonds – with varying coupons, maturities, embedded options, and covenant stipulations – has historically posed substantial challenges to data collection and maintenance, limiting the scope and depth of empirical studies.

Not surprisingly, existing research on corporate bonds relies on datasets with relatively short time series, typically starting in 2002. This limited sample period inevitably constrains the power of empirical tests and excludes critical historical periods. Moreover, several high-profile studies have faced strong criticism due to data and replicability issues (see Dickerson, Mueller, and Robotti (2023) for more information). Despite past efforts, significant gaps persist in our understanding of corporate bond pricing, underscoring the need for more comprehensive data to advance research in this area.

This paper addresses these limitations by constructing a new historical corporate bond database spanning 128 years, from 1895 to 2022. The database integrates hand-collected monthly bond quotes from three key archival print sources – *Commercial and Financial Chronicle*, *Standard & Poor's Bond Guide*, and *Moody/Mergent's Bond Record* – with existing datasets such as Lehman Brothers and Wharton Research Data Services (WRDS). By combining these sources, we create monthly cross-sections of corporate bond records, offering unparalleled coverage and granularity that are unattainable from any single source. The resulting database encompasses nearly 110,000 unique bonds and 8 million observations.

The database construction process involves meticulous data collection and validation, including digitizing over 80,000 pages of archival material and implementing rigorous double-blind data entry protocols to ensure accuracy. Consolidating bond quotes from different historical sources also requires careful attention, as common characteristics of bonds, market practices, and reporting conventions have evolved over time. We confirm that the resulting database demonstrates remarkable consistency across sources, with overlapping periods showing strong agreement both in coverage (bond counts) and pricing (bond yields).

To enhance our database further, we integrate bond characteristics from *Moody's Manual*, a comprehensive reference series that contains detailed information on companies and their securities. First published in 1900, this manual provides data for nearly all historical bonds in our extensive sample. Drawing from over half a million manually scanned pages, Moody's Manual fills a critical gap in bond characteristics that standard datasets like Mergent FISD (Fixed Income Securities Database), which focus on more recent periods, cannot fill. Taking a further step, we merge our corporate bond database with the CRSP (Center for Research in Security Prices) stock database, making it possible to examine the joint return dynamics between stocks and corporate bonds issued by the same firms.

Leveraging this extensive time series and large cross-section, we explore a range of fundamental asset pricing topics. We begin by converting corporate bond prices into yields and analyzing their historical time series from 1895. The median yield time series reveals elevated levels during the Great Depression and a surge in the inflationary period of the 1970-80s. To isolate the credit component, we compute credit spreads by subtracting Treasury yields of matching maturity from corporate bond yields. The median credit spread time series shows two notable peaks: 8% during the Great Depression and 6% during the 2008 global financial crisis. In addition, we examine the GZ spread introduced by Gilchrist and Zakrajšek (2012), which represents an equally-weighted average of credit spreads for senior unsecured bonds issued by U.S. public non-financial firms. Using our extended database, we significantly lengthen the time series of the GZ spread, which is known to contain valuable information about future economic activity. Our analysis validates the economic content of the GZ spread over a much longer sample with more recessions and extreme business cycle fluctuations, while revealing that its predictive power varies across different historical episodes.

Next, we compute returns on individual corporate bonds, which we aggregate to construct the bond market factor. Equipped with the time series of the bond market factor over a much longer sample than previously available, we revisit recent studies on which risks are priced in the corporate bond market. In a comprehensive analysis, Dickerson, Mueller, and Robotti (2023) document striking empirical evidence that neither well-known bond factors nor prominent stock factors explain the cross-section of corporate bond returns beyond the bond market factor. This suggests a puzzling disconnect not only between the stock and bond markets but also between theory and empirics: If a given factor captures systematic risk to the extent that it generates significant risk premia in the cross-section of stock returns, why isn't it priced in corporate bonds?

Our analysis reveals that these conclusions may reflect the limited statistical power of short samples typically used in corporate bond research, rather than true economic patterns. When we extend the sample back to 1926, various risk factors emerge as significant sources of cross-sectional pricing. Beyond the now-strongly-significant bond market factor, not only the downside and credit risk factors of Bai, Bali, and Wen (2019) but also traditional stock factors such as the stock market, size, and value factors of Fama and French (1993) add incremental explanatory power and remain statistically significant. This pattern extends to nontraded macroeconomic factors such as the intermediary factor of He, Kelly, and Manela (2017), macro uncertainty factor of Jurado, Ludvigson, and Ng (2015), and tail risk factor of Kelly and Jiang (2014), all priced with theoretically consistent signs.

While it is not our intention to seek the best factor model, the evidence from our long sample brings some clarity to the pricing of corporate bonds. The cross-section of corporate bond returns appears to share a common factor structure with that of stock returns; the same risk factors are both priced and commanding meaningful risk premia across the two markets. We believe that our research can serve as an important stepping stone for better understanding the joint factor structure between the two markets.

Overall, the historical corporate bond database we construct addresses a critical shortcoming in the literature by offering long-term perspectives on corporate bond prices and returns. The richness of the data supports more rigorous empirical testing of market dynamics for corporate bonds, overcoming the limitations of existing short-sample datasets. Our work demonstrates that we can construct a "CRSP" for corporate bonds by exploiting archival print sources, establishing a new benchmark for corporate bond research.

Related literature

Our paper contributes to the extensive body of research on long-run asset returns, focusing on the perspective of corporate bond investors. Dimson, Marsh, and Staunton (2009) conduct one of the first broad analysis of long-term investment returns across multiple countries. More recently, Jordà, Schularick, and Taylor (2019) provide a comprehensive examination of asset returns over nearly 150 years for stocks, housing, government bonds. However, their analysis does not include corporate bonds. Chambers, Dimson, Ilmanen, and Rintamäki (2024) highlight the significant challenges in studying historical corporate bond markets, citing the sparse availability of data on quotes, credit ratings, defaults, and recovery rates; they point out that these limitations make it difficult to produce reliable and consistent estimates of credit risk premia before 1973. We address these issues with our corporate bond database, providing a robust foundation for analyzing the long-run performance of corporate bonds.[1]

Equipped with our corporate bond database, we also contribute to the literature on credit cycles, which examines how changes in credit market conditions influence real economic activity. Gilchrist and Zakrajšek (2012) investigate the role of fluctuations in a newly constructed measure of corporate credit spreads, the so-called GZ spread, as predictors of recessions. They also calculate the excess bond premium, a measure that isolates credit market sentiment from expected default risk, using firm-level data on corporate bond prices. Related, Greenwood and Hanson (2013) show that during credit booms, the credit quality of corporate debt issuers deteriorates, signaling an overheating of credit markets, which, in turn, predicts lower excess returns for corporate bondholders. Building on this work, López-Salido, Stein, and Zakrajšek (2017) demonstrate that elevated credit-market sentiment, measured through the spread between Baa-rated industrial bond yields and Treasury yields, predicts a decline in

[1] For early examples of research on asset returns over an extended period of time, see Ibbotson and Sinquefield (1976), Siegel (1992), and Siegel (1994). For a recent analysis of historical corporate bond returns on the Brussels Stock Exchange, see Van Mencxel, Annaert, and Deloof (2024).

future economic activity. Our study adds to this body of literature by providing an in-depth analysis of various credit spread measures over an extended sample period, offering new insights into their role in forecasting economic fluctuations.

Lastly, our work closely relates to the growing literature on the cross-section of corporate bond returns. Since the introduction of the two-factor model by Fama and French (1993), substantial progress has been made in identifying and evaluating factor structures within the bond market. Numerous stock market factors have been adapted and studied in the context of bond returns.[2] Recent studies, including Dickerson, Mueller, and Robotti (2023), Dick-Nielsen, Feldhütter, Pedersen, and Stolborg (2024), and Jostova, Nikolova, and Philipov (2024), highlight significant replication challenges, emphasizing the need for standardized methodologies for data cleaning, outlier and error handling, factor construction, and inference. We attempt to mitigate these concerns with our extensive database, offering a robust resource for researchers. By significantly extending the data sample period for corporate bonds, we provide new perspectives on the factor structure of corporate bond returns and contribute to broader asset pricing topics.

The rest of the paper proceeds as follows. Section 2 describes the construction of our historical corporate bond database, detailing the hand-collection process, data validation, and integration with existing datasets. Section 3 examines the historical time series of corporate bond yields, highlighting key trends and new insights into the evolution of credit spreads and the GZ spread over time. Section 4 calculates excess returns for individual bonds and constructs a corporate bond market factor, which we test and find to be priced in the cross-section. Section 5 examines whether prominent bond and stock factors, as well as nontraded factors, are priced beyond the bond market factor over our extended sample period. Finally, Section 6 concludes.

[2] See, for instance, Gebhardt, Hvidkjaer, and Swaminathan (2005), Lin, Wang, and Wu (2011), Jostova, Nikolova, Philipov, and Stahel (2013), Chordia, Goyal, Nozawa, Subrahmanyam, and Tong (2017), Choi and Kim (2018), Bai, Bali, and Wen (2019), Chung, Wang, and Wu (2019), Bali, Subrahmanyam, and Wen (2021), Kelly, Palhares, and Pruitt (2023), Elkamhi, Jo, and Nozawa (2024), and van Binsbergen, Nozawa, and Schwert (2024).

2 A new historical corporate bond database

To construct a historical corporate bond database that spans over a century, we complement existing datasets by hand-collecting monthly corporate bond quotes from three separate sources: the (i) Commercial and Financial Chronicle, (ii) Standard & Poor's Bond Guide, and (iii) Mergent/Moody's Bond Record. Figure 1 illustrates the data coverage periods for these three new sources as well as for two existing datasets from Lehman Brothers and Wharton Research Data Services (WRDS). By combining these, we create monthly cross-sections of corporate bond records spanning 128 calendar years, from 1895 to 2022.

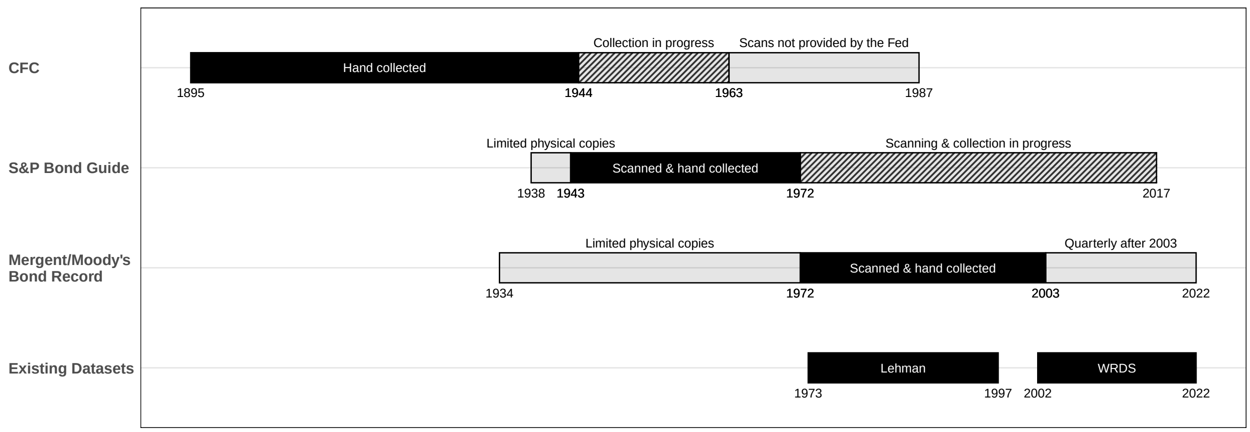


Figure 1: Coverage of the new corporate bond database. This figure illustrates the construction of our historical corporate bond database, showing data coverage periods for five sources: (i) Commercial and Financial Chronicle (CFC), (ii) Standard & Poor's (S&P) Bond Guide, and (iii) Mergent/Moody's Bond Record, (iv) Lehman Brothers, and (v) Wharton Research Data Services (WRDS). The first three sources are hand-collected and represent novel additions, while the latter two come from existing datasets. The bars represent the availability of each data source, with darker shading indicating periods for which we collect data. Annotations provide explanations for uncollected data periods. The resulting database comprises monthly cross-sections of corporate bond records spanning 128 calendar years, from April 1895 to June 2022.

In this section, we detail the construction of our historical corporate bond database. We begin by describing the process of hand-collecting monthly corporate bond records (Section 2.1). We then explain how we enhance the data by adding comprehensive bond characteristics, manually collected from Moody's Manual (Section 2.2). Following this, we outline our methodology for aggregating the data to construct price time series (Section 2.3). Finally, we discuss the integration of our corporate bond database with CRSP (Center for Research in Security Prices), which provides stock prices on all publicly traded firms dating back to 1926 (Section 2.4).

2.1 Hand-collecting monthly corporate bond records

2.1.1 Commercial and Financial Chronicle: 1895-1963

The Commercial and Financial Chronicle (in short, CFC), a weekly business newspaper published from 1865 to 1987, provides comprehensive information on the industrial and commercial sectors. Although published weekly, the CFC offers bond records at monthly intervals through supplementary publications.[3] Scanned digital records of the CFC are obtained from the Federal Reserve Archival System for Economic Research (FRASER), a digital archive maintained by the Federal Reserve Bank of St. Louis. Our monthly corporate bond data from the CFC spans May 1895 to December 1963, corresponding to the period for which FRASER provides digital records.[4]

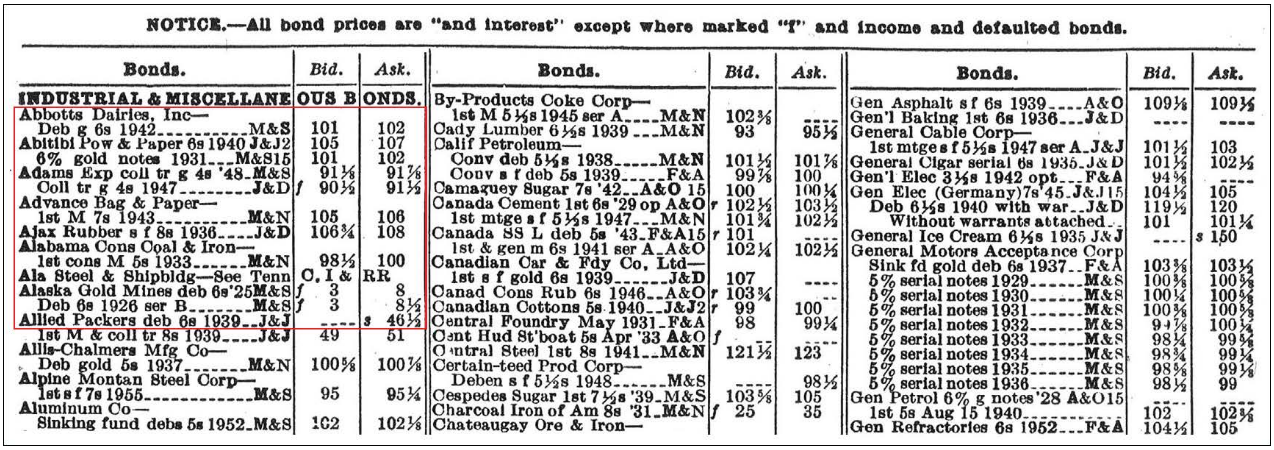
Within these publications, bond quotes are provided under "General Quotations." We collect bond records for "Railroad" and "Miscellaneous Securities" from May 1895 until August 1916. This categorization reflects that railroad bonds were a dominant part of the early stage U.S. corporate bond market. Railroads were among the largest infrastructure projects of the time, requiring massive long-term capital investments, largely financed through bonds. As the market evolved and expanded toward other industries, this categorization scheme was updated in August 1916 to include "Railroad," "Public Utilities," and "Industrial and Miscellaneous Securities," under which we continue our collection.

We convert these scanned records, totaling over 15,000 pages, into a machine-readable format. To illustrate, Panel A of Figure 2 presents an example of a scanned original copy from the CFC (specifically, a partial segment from the March 1928 issue of the "Bank and Quotation Record"). This excerpt displays bond records under "Industrial and Miscellaneous Securities," which encompasses all issued securities not classified under "Railroad" or "Public Utilities." From this sample segment, we can observe that the CFC provides bid and ask

[3]These records appear in the "Quotation Supplement" from May 1895 to February 1902, the "Bank and Quotation Supplement/Section" from March 1902 to February 1928, and the "Bank and Quotation Record" from March 1928 to March 1987.

[4]Even within this period, some records are missing in FRASER, such as the January and February 1928 issues. In such cases, we obtain physical copies and manually scan them. So far, we have completed the data collection up to December 1944.

Panel A: Scanned original



NOTICE.—All bond prices are "and interest" except where marked "f" and income and defaulted bonds.

Bonds.	Bid.	Ask.
INDUSTRIAL & MISCELLANEOUS BONDS.		
Abbotts Dairies, Inc—		
Deb g 6s 1942..........M&S	101	102
Abitibi Pow & Paper 6s 1940 J&J2	105	107
6% gold notes 1931...M&S15	101	102
Adams Exp coll tr g 4s '48..M&S	91⅛	91⅞
Coll tr g 4s 1947.........J&D	f 90½	91½
Advance Bag & Paper—		
1st M 7s 1943..........M&N	105	106
Ajax Rubber s f 8s 1936....J&D	106¾	108
Alabama Cons Coal & Iron—		
1st cons M 5s 1933......M&N	98½	100
Ala Steel & Shipbldg—See Tenn	C, I &	RR
Alaska Gold Mines deb 6s'25M&S	f 3	8
Deb 6s 1926 ser B........M&S	f 3	8½
Allied Packers deb 6s 1939..J&J	----	s 46½
1st M & coll tr 8s 1939....J&J	49	51
Allis-Chalmers Mfg Co—		
Deb gold 5s 1937.......M&N	100⅝	100⅞
Alpine Montan Steel Corp—		
1st s f 7s 1955.............M&S	95	95¼
Aluminum Co—		
Sinking fund debs 5s 1952..M&S	102	102⅛

Bonds.	Bid.	Ask.
By-Products Coke Corp—		
1st M 5½s 1945 ser A.....M&N	102⅜	----
Cady Lumber 6½s 1939 ...M&N	93	95½
Calif Petroleum—		
Conv deb 5½s 1938......M&N	101½	101⅞
Conv s f deb 5s 1939......F&A	99⅞	100
Camaguey Sugar 7s '42..A&O 15	100	100¼
Canada Cement 1st 6s '29 op A&O	r 102½	103½
1st mtge s f 5½s 1947...M&N	101¾	102½
Canada SS L deb 5s '43..F&A15	r 101	----
1st & gen m 6s 1941 ser A..A&O	102¼	102½
Canadian Car & Fdy Co, Ltd—		
1st s f gold 6s 1939.......J&D	107	----
Canad Cons Rub 6s 1946..A&O	r 103¾	--
Canadian Cottons 5s 1940..J&J2	r 99	100
Central Foundry May 1931..F&A	98	99¼
Cent Hud St'boat 5s Apr '33 A&O	f --	----
Central Steel 1st 8s 1941..M&N	121½	123
Certain-teed Prod Corp—		
Deben s f 5½s 1948.......M&S	----	98½
Cespedes Sugar 1st 7½s '39..M&S	103⅝	105
Charcoal Iron of Am 8s '31..M&N	f 25	35
Chateaugay Ore & Iron—		

Bonds.	Bid.	Ask.
Gen Asphalt s f 6s 1939.....A&O	109⅛	109½
Gen'l Baking 1st 6s 1936...J&D	----	----
General Cable Corp—		
1st mtge s f 5½s 1947 ser A..J&J	101½	103
General Cigar serial 6s 1935..J&D	101½	102½
Gen'l Elec 3½s 1942 opt....F&A	94⅝	----
Gen Elec (Germany) 7s'45..J&J15	104½	105
Deb 6½s 1940 with war...J&D	119½	120
Without warrants attached...	101	101¼
General Ice Cream 6½s 1935 J&J	----	s 150
General Motors Acceptance Corp		
Sink fd gold deb 6s 1937..F&A	103⅜	103½
5% serial notes 1929.......M&S	100⅜	100⅝
5% serial notes 1930.......M&S	100¼	100⅝
5% serial notes 1931.......M&S	100⅜	100⅝
5% serial notes 1932.......M&S	99⅞	100¼
5% serial notes 1933.......M&S	98¼	99⅝
5% serial notes 1934.......M&S	98¾	99¼
5% serial notes 1935.......M&S	98⅝	99⅛
5% serial notes 1936.......M&S	98½	99
Gen Petrol 6% g notes '28 A&O15	----	----
1st 5s Aug 15 1940............	102	102⅜
Gen Refractories 6s 1952....F&A	104½	105

Panel B: Data entry result

		Quotes								
	Coupon	Freelancer A			Freelancer B			Freelancer C		
Company/Bond Name	Dates	Bid	Ask	Prefix	Bid	Ask	Prefix	Bid	Ask	Prefix
Abbotts Dairies g deb 6s 1942	M&**S**	101	102		101	102		101	102	
Abitibi Pow & Paper 6s 1940	**J**&J 12	105	107		105	107		105	107	
Abitibi Pow & Paper 6%...	M&S 15	101	102		101	102		101	102	
Adams Exp coll tr g 4s '48	M&S	91 1/8	91 7/8		91 1/8	91 7/8		91 1/8	91 7/8	
Adams Exp coll tr g 4s 1947	**J**&D	90 1/2	91 1/2	*f*	90 1/2	91 1/2	*f*	90 1/2	91 1/2	*f*
Advance Bag & Paper 1st M...	M&N	105	106		105	108		105	**106**	
Ajax Rubber s f 8s 1936	J&**D**	106 3/4	108		106 3/4	108		106 3/4	108	
Alabama Cons Coal & Iron...	M&N	98 1/2	100		98 1/2	100		98 1/2	100	
Alaska Gold Mines deb 6s '25	M&S	3	8	*f*	3	8	*f*	3	8	*f*
Alaska Gold Mines deb 6s...	M&S	3	8 1/2	*f*	3	8 1/2	*f*	3	8 1/2	*f*
Allied Packers deb 6s 1939	J&**J**		46 1/2	*s*		46 1/2	*s*		46 1/2	*s*
⋮										

Figure 2: Commercial and Financial Chronicle. This figure illustrates our data collection process from the Commercial and Financial Chronicle. Panel A shows a scanned segment from the May 1926 issue of the "Bank and Quotation Section." Panel B displays the results of the data entry process, which involves two independent freelancers separately entering each data item, followed by a third freelancer comparing these entries. When no discrepancy is detected, the entry is accepted. In cases of discrepancy, the third freelancer consults the original document to make a final determination, effectively breaking the tie.

prices, along with detailed information for each bond, including company/bond name and coupon dates. By default, prices are reported as clean prices per \$100.[5] Hence, the actual amount paid when trading a bond (i.e., dirty price) should be calculated by adding accrued interest to the reported price, unless the prefix "*f*" (flat) is attached. Note that defaulted bonds trade flat without accrued interest because the issuer no longer pays interest. Similarly, income bonds (more prevalent in the past, notably issued by railroad companies under financial hardship) also trade flat as the issuer is not obligated to pay interest unless it has sufficient

[5]The CFC adopted this convention from the February 1909 issue, following the New York Stock Exchange's transition to quoting all bonds in terms of clean prices starting January 2, 1909. Before this change, the CFC primarily reported dirty prices and used a specific mark to indicate when clean prices were reported.

earnings to do so.[6] We can see other prefixes in Panel A: "*s*" denotes a sale price, and "*r*" indicates a Canadian dollar price. Other abbreviations, symbols, and prefixes used by the CFC are reported in the Internet Appendix.

Panel B shows the results of the data entry process. The first column contains company/bond name information, which is later parsed to extract embedded details. For instance, the first row entry "Abbotts Dairies Deb g 6s 1942" refers to a debenture ("Deb") issued by Abbotts Dairies, Inc. with a 6% coupon rate and a 1942 maturity year. The symbol "g" indicates that the bond embedded a gold clause, a historical provision that guaranteed interest and principal payments in gold or its equivalent value to protect investors from currency devaluation.[7] The second column specifies semi-annual coupon dates. The letter pairing notation describes the months of these payments, with the bold letter marking the month of the last coupon payment at maturity. A number following the letter pair, if any, indicates the payment date within the month; absence of a number implies payments occur on the first day of the month. For example, "**M**&S 15" in the third row denotes semi-annual interest payments on the 15th of March and September, with the last coupon payment made in March. We cross-check these bond characteristics using Moody's Manual as later described in Section 2.2.

These columns are followed by the data entries of bid, ask, and sale prices, which are our main objects of interest. For accuracy, we implement a rigorous double-blind data entry process involving three independent freelancers. Freelancers A and B separately enter each data item from the scanned original, while Freelancer C compares these entries. The case of "Advance Bag & Paper" illustrates our resolution process for discrepancies. Here, Freelancers A and B disagree on the ask price (106 vs. 108; marked in red). Freelancer C resolves the discrepancy by consulting the original document, confirming that Freelancer A's entry is correct. The accuracy of individual data entry freelancers typically exceeds 99%. This high

[6]The CFC often puts "*f*" to non-defaulted/income bonds. This implies that the reported price is a dirty price that already includes accrued interest.

[7]Gold clauses were unilaterally annulled by Congress in June 1933. See Edwards (2018) for the historical background behind this abrogation, Edwards, Longstaff, and Marin (2015) for its effects on sovereign debt markets, and Gomes, Kilic, and Plante (2020) for the impact on firm investment surrounding legal challenges to its constitutionality.

level of accuracy, combined with our double-blind data entry process, should result in data that closely reflect the original documents.

2.1.2 Standard & Poor's Bond Guide: 1943-1972

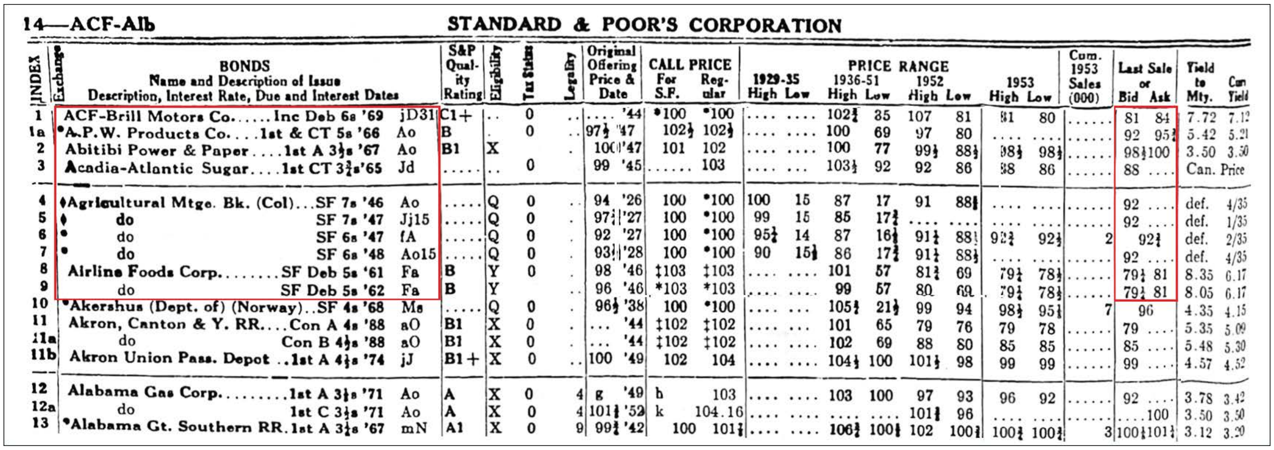
The Standard & Poor's Bond Guide (hereafter, S&P Bond Guide) provides comprehensive information on corporate bonds and their prices from 1938 to 2017.[8] To extract monthly corporate bond quotes from the S&P Bond Guide, we first collect the physical copies, which we then manually scan. Our data coverage is only partial, from January 1943 to December 1972, due to two reasons. First, locating earlier issues is particularly challenging.[9] Second, although we have managed to complete most of the scans for the post-1972 period, data collection is still in progress. We assign a relatively lower priority on extending the sample as this period is fully covered by other data sources. Nevertheless, we plan to continue our data collection for robustness checks. The current sample corresponds to over 50,000 pages of scanned records.

Panel A of Figure 3 shows an excerpt from a scanned original copy of the S&P Bond Guide, specifically a partial segment from the March 1953 issue. Due to space limitations, only half of the information provided for these bonds is shown; the complete record spans two pages, and the second page is available in the Internet Appendix. Similar to the CFC, the company/bond name, coupon dates, and monthly bond quotes are provided. In the "Last Sale or Bid Ask" column, bid and ask prices are represented by two numbers, while a single number centered in the column indicates a sale price. All prices are clean, and accrued interest should be added for transactions, unless noted with "flat." On top of the pricing data, the S&P Bond Guide offers additional bond characteristics, such as the S&P rating (third column), yield to maturity/current yield (last column), and amount outstanding (see the Internet Appendix). The term "def" in the yields column signifies that the bond was defaulted in its payment since the date reported next to it. The label "Can. Price" illustrates that the bond is denominated

[8] Prior to May 1941, when Standard Statistics merged with Poor's Publishing Company, this publication was known as Poor's Bond Guide.

[9] Related, we are unable to obtain the October 1945 and March 1948 issues.

Panel A: Scanned original



14—ACF-Alb STANDARD & POOR'S CORPORATION

INDEX	BONDS Name and Description of Issue Description, Interest Rate, Due and Interest Dates		S&P Quality Rating	Eligibility	Tax Status	Legality	Original Offering Price & Date	Call Price For S.F.	Call Price Regular	1929-35 High	1929-35 Low	1936-51 High	1936-51 Low	1952 High	1952 Low	1953 High	1953 Low	Cum. 1953 Sales (000)	Last Sale or Bid	Ask	Yield to Mty.	Cur Yield
1	ACF-Brill Motors Co......Inc Deb 6s '69	jD31	C1+	..	0	..	.... '44	•100	•100	....	....	102¾	35	107	81	81	80	......	81	84	7.72	7.12
1a	•A.P.W. Products Co....1st & CT 5s '66	Ao	B	.	0	..	97½ '47	102½	102½	....	....	100	69	97	80	....	....	......	92	95¾	5.42	5.21
2	Abitibi Power & Paper....1st A 3½s '67	Ao	B1	X		.	100 '47	101	102	....	....	100	77	99½	88½	98½	98½	......	98½	100	3.50	3.50
3	Acadia-Atlantic Sugar....1st CT 3¾s'65	Jd	.....	..	0		99 '45	......	103	....	...	103½	92	92	86	88	86	......	88	....	Can. Price	
4	♦Agricultural Mtge. Bk. (Col)...SF 7s '46	Ao	.....	Q	0	..	94 '26	100	•100	100	15	87	17	91	88⅝	....	....	......	92	....	def.	4/35
5	♦ do SF 7s '47	Jj15	.....	Q	0	.	97¾ '27	100	•100	99	15	85	17¾	....	....	....	....	......	92	....	def.	1/35
6	• do SF 6s '47	fA	.....	Q	0	.	92 '27	100	•100	95¾	14	87	16⅛	91¼	88½	92¾	92½	2	92¾		def.	2/35
7	• do SF 6s '48	Ao15	.....	Q	0	.	93¾ '28	100	•100	90	15⅛	86	17¾	91¼	88½	....	....	......	92	....	def.	4/35
8	Airline Foods Corp........SF Deb 5s '61	Fa	B	Y	0	.	98 '46	‡103	‡103	....	....	101	57	81¾	69	79¼	78½	......	79¼	81	8.35	6.17
9	do SF Deb 5s '62	Fa	B	Y		..	96 '46	*103	*103	....	....	99	57	80	69	79¼	78½	......	79¼	81	8.05	6.17
10	•Akershus (Dept. of) (Norway)..SF 4s '68	Ms	.....	Q	0	.	96½ '38	100	•100	....	...	105½	21½	99	94	98½	95⅛	7	96		4.35	4.15
11	Akron, Canton & Y. RR....Con A 4s '88	aO	B1	X	0	.	... '44	‡102	‡102	....	...	101	65	79	76	79	78	......	79	....	5.35	5.09
11a	do Con B 4½s '88	aO	B1	X	0	.	... '44	‡102	‡102	....	....	102	69	88	80	85	85	......	85	....	5.48	5.30
11b	Akron Union Pass. Depot ..1st A 4½s '74	jJ	B1+	X	0	..	100 '49	102	104	....	....	104½	100	101½	98	99	99	......	99	....	4.57	4.52
12	Alabama Gas Corp.........1st A 3½s '71	Ao	A	X	0	4	g '49	h	103	....	....	103	100	97	93	96	92	......	92	....	3.78	3.42
12a	do 1st C 3½s '71	Ao	A	X	0	4	101⅛ '52	k	104.16	....	....	....	....	101⅛	96	....	....	......	....	100	3.50	3.50
13	•Alabama Gt. Southern RR.1st A 3¼s '67	mN	A1	X	0	9	99¾ '42	100	101¾	....	....	106¾	100⅛	102	100⅛	100¾	100¾	3	100⅛	101½	3.12	3.20

Panel B: Data entry result

		Quotes								
	Coupon	Freelancer A			Freelancer B			Freelancer C		
Company/Bond Name	Dates	Bid	Ask	Sale	Bid	Ask	Sale	Bid	Ask	Sale
ACF-Brill Motors Co inc...	jD 31	81	84		81	84		81	84	
A.P.W. Products Co. 1st &...	Ao	92	95 3/4		92	95 3/4		92	95 3/4	
Abitibi Power & Paper 1st...	Ao	98 1/2	100		98 1/2	100		98 1/2	100	
Acadia-Atlantic Sugar 1st...	Jd	88			88			88		
Agricultural Mtge Bk. (Col)...	Ao	92			92			92		
Agricultural Mtge Bk. (Col)...	Jj 15	92			92			92		
Agricultural Mtge Bk. (Col)...	fA			92 3/4			92 3/4			92 3/4
Agricultural Mtge Bk. (Col)...	Ao 15	92			92			92		
Airline Foods Corp. sf deb...	Fa	79 1/4	81		79 1/4	81		79 1/4	81	
Airline Foods Corp. sf deb...	Fa	79 1/4	81		79 1/4	81		79 1/4	81	
⋮										

Figure 3: Standard & Poor's Bond Guide. This figure illustrates our data collection process from the Standard & Poor's Bond Guide. Panel A shows a scanned segment from the March 1953 issue. Panel B displays the results of the data entry process, which involves two independent freelancers separately entering each data item, followed by a third freelancer comparing these entries. When no discrepancy is detected, the entry is accepted. In cases of discrepancy, the third freelancer consults the original document to make a final determination, effectively breaking the tie.

in Canadian dollars.

Panel B of the figure demonstrates our data entry process for these scanned bond records. As with the CFC, the first column contains company/bond name. The coupon dates column shows when semi-annual interest payments are made to bondholders, with the month of the last coupon at maturity denoted by a capital letter. Unless a date is specified, interest is paid on the first day of the month. As previously described, the data entry process for bid, ask, and sale prices is double-blind, involving three freelancers: Freelancers A and B independently enter the data, and Freelancer C serves as the final reviewer to resolve any discrepancies.

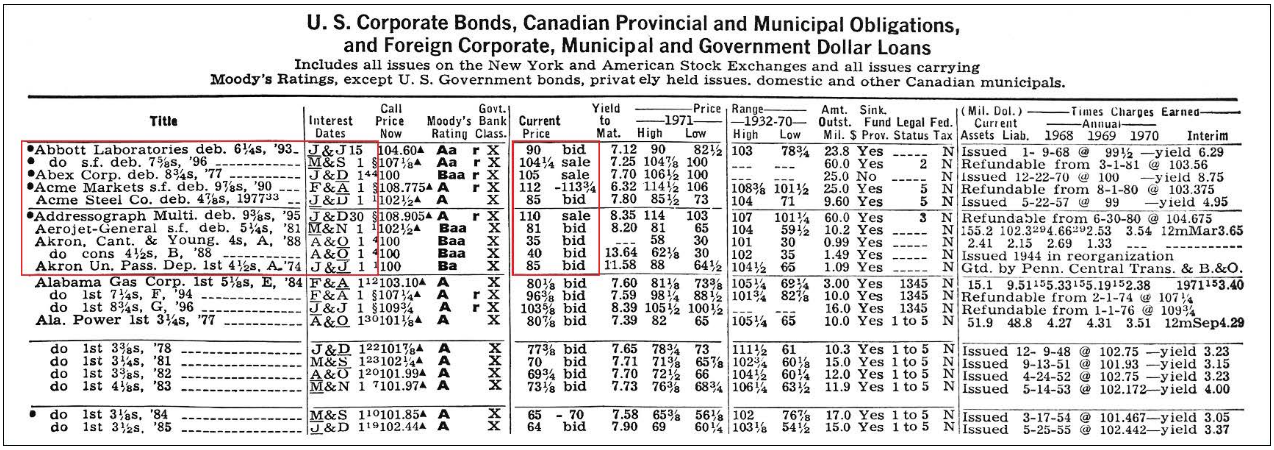
2.1.3 Mergent/Moody's Bond Record: 1972-2003

Moody's Bond Record, a print publication first released in 1934, has long been a key resource for bond market information. In 1998, Moody's Investors Service sold its publishing division to Financial Communications, which later became Mergent, leading to the rebranding of the publication as Mergent's Bond Record in July 1999. We extract monthly corporate bond data from this data source by collecting and manually scanning physical copies for the period between January 1972 and December 2003. We begin our data collection from 1972, although the publication dates back to 1934, due to difficulties in finding earlier physical records. We stop our data collection at 2003, despite its continued availability beyond 2003, because its publication frequency changed to quarterly thereafter. During our sample period, we scan and collect a total of 15,000 pages by hand.

Panel A of Figure 4 displays a scanned original copy of Mergent/Moody's Bond Record, featuring a portion of the January 1972 issue. Like the CFC and the S&P Bond Guide, it includes information such as the company/bond name, coupon dates, and monthly bond quotes. While more condensed than the S&P Bond Guide, it contains comparable details, including Moody's Rating, yield to maturity, and the amount outstanding. The letter "*r*" next to Moody's rating indicates that bonds are issued in fully registered form rather than bearer form. The "Current Price" column typically contains bids or sales; if a single price is listed, an accompanying note describes the type (bid/sale), but if there are two numbers, they represent bid and ask prices. All prices are clean. While not appearing on this excerpt, the price is marked with the letter "*f*" when the bond trades flat without accrued interest. Interest or principal default is indicated in the "Interest Dates" column with symbols † and ‡, respectively, along with the date of default. The Internet Appendix provides a different page of the January 1972 issue, which contains such examples.

Panel B highlights the data entry process. The first column contains company/bond name. The coupon dates column shows when semi-annual interest payments are made, with an underlined letter denoting the final month of maturity and a number showing the day of the month. As outlined earlier, bid, ask, and sale prices are entered using a double-blind

Panel A: Scanned original



U. S. Corporate Bonds, Canadian Provincial and Municipal Obligations, and Foreign Corporate, Municipal and Government Dollar Loans

Includes all issues on the New York and American Stock Exchanges and all issues carrying Moody's Ratings, except U. S. Government bonds, privat ely held issues. domestic and other Canadian municipals.

Title	Interest Dates	Call Price Now	Moody's Rating	Govt. Bank Class.	Current Price	Yield to Mat.	Price 1971 High	Price 1971 Low	Range 1932-70 High	Range 1932-70 Low	Amt. Outst. Mil. $	Sink. Fund Prov.	Legal Status	Fed. Tax	(Mil. Dol.) Current Assets Liab. — Times Charges Earned — Annual 1968 1969 1970 Interim
•Abbott Laboratories deb. 6¼s, '93	J&J 15	104.60▲	Aa	r X	90 bid	7.12	90	82½	103	78¾	23.8	Yes	-----	N	Issued 1- 9-68 @ 99½ —yield 6.29
• do s.f. deb. 7⅝s, '96	M&S 1	§107⅛▲	Aa	r X	104¼ sale	7.25	104⅞	100	---	---	60.0	Yes	2	N	Refundable from 3-1-81 @ 103.56
•Abex Corp. deb. 8¾s, '77	J&D 1[44]	100	Baa	r X	105 sale	7.70	106½	100	---	---	25.0	No	-----	N	Issued 12-22-70 @ 100 —yield 8.75
•Acme Markets s.f. deb. 9⅞s, '90	F&A 1	§108.775▲	A	r X	112 -113¾	6.32	114½	106	108⅜	101½	25.0	Yes	5	N	Refundable from 8-1-80 @ 103.375
Acme Steel Co. deb. 4⅞s, 1977[33]	J&D 1	[1]102½▲	A	X	85 bid	7.80	85½	73	104	71	9.60	Yes	5	N	Issued 5-22-57 @ 99 —yield 4.95
•Addressograph Multi. deb. 9⅜s, '95	J&D30	§108.905▲	A	r X	110 sale	8.35	114	103	107	101¼	60.0	Yes	3	N	Refundable from 6-30-80 @ 104.675
Aerojet-General s.f. deb. 5¼s, '81	M&N 1	[1]102½▲	Baa	X	81 bid	8.20	81	65	104	59½	10.2	Yes	-----	N	155.2 102.3[29]4.66[29]2.53 3.54 12mMar3.65
Akron, Cant. & Young. 4s, A, '88	A&O 1	[4]100	Baa	X	35 bid	---	58	30	101	30	0.99	Yes	-----	N	2.41 2.15 2.69 1.33 --- ----------
do cons 4½s, B, '88	A&O 1	[4]100	Baa	X	40 bid	13.64	62⅛	30	102	35	1.49	Yes	-----	N	Issued 1944 in reorganization
Akron Un. Pass. Dep. 1st 4½s, A '74	J&J 1	[1]100	Ba	X	85 bid	11.58	88	64½	104½	65	1.09	Yes	-----	N	Gtd. by Penn. Central Trans. & B.&O.
Alabama Gas Corp. 1st 5⅛s, E, '84	F&A 1[12]	103.10▲	A	X	80⅛ bid	7.60	81⅛	73⅜	105¼	69¼	3.00	Yes	1345	N	15.1 9.51[15]5.33[15]5.19[15]2.38 1971[15]3.40
do 1st 7¼s, F, '94	F&A 1	§107¼▲	A	r X	96⅜ bid	7.59	98¼	88½	101¾	82⅞	10.0	Yes	1345	N	Refundable from 2-1-74 @ 107¼
do 1st 8¾s, G, '96	J&J 1	§109¾	A	r X	103⅝ bid	8.39	105½	100½	---	---	16.0	Yes	1345	N	Refundable from 1-1-76 @ 109¾
Ala. Power 1st 3¼s, '77	A&O 1[30]	101⅛▲	A	X	80⅞ bid	7.39	82	65	105¼	65	10.0	Yes	1 to 5	N	51.9 48.8 4.27 4.31 3.51 12mSep4.29
do 1st 3⅜s, '78	J&D 1[22]	101⅞▲	A	X	77⅜ bid	7.65	78¾	73	111½	61	10.3	Yes	1 to 5	N	Issued 12- 9-48 @ 102.75 —yield 3.23
do 1st 3¼s, '81	M&S 1[23]	102¼▲	A	X	70 bid	7.71	71⅜	65⅞	102¾	60⅛	15.0	Yes	1 to 5	N	Issued 9-13-51 @ 101.93 —yield 3.15
do 1st 3⅜s, '82	A&O 1[20]	101.99▲	A	X	69¾ bid	7.70	72½	66	104½	60¼	12.0	Yes	1 to 5	N	Issued 4-24-52 @ 102.75 —yield 3.23
do 1st 4⅛s, '83	M&N 1	[7]101.97▲	A	X	73⅛ bid	7.73	76⅜	68¾	106¼	63½	11.9	Yes	1 to 5	N	Issued 5-14-53 @ 102.172—yield 4.00
• do 1st 3⅛s, '84	M&S 1[10]	101.85▲	A	X	65 - 70	7.58	65⅜	56⅛	102	76⅞	17.0	Yes	1 to 5	N	Issued 3-17-54 @ 101.467—yield 3.05
do 1st 3½s, '85	J&D 1[19]	102.44▲	A	X	64 bid	7.90	69	60¼	103⅛	54½	15.0	Yes	1 to 5	N	Issued 5-25-55 @ 102.442—yield 3.37

Panel B: Data entry result

Company/Bond Name	Coupon Dates	Freelancer A Bid	Freelancer A Ask	Freelancer A Sale	Freelancer B Bid	Freelancer B Ask	Freelancer B Sale	Freelancer M Bid	Freelancer M Ask	Freelancer M Sale
Abbott Laboratories deb...	J&J 15	90			90			90		
Abbott Laboratories s.f. deb...	M&S 1			104.25			104.25			104.25
Abex Corp. deb. 8 3/4s, '77	J&D 1			105			105			105
Acme Markets s.f. deb. 9...	F&A 1	112	113.75		112	113.75		112	113.75	
Acme Steel Co. deb. 4 7/8s...	J&D 1	85			85			85		
Addressograph Multi. deb...	J&D 30			110			110			110
Aerojet-General s.f. deb. 5...	M&N 1	81			81			81		
Akron, Cant. & Young. 4s...	A&O 1	35			35			35		
Akron, Cant. & Young...	A&O 1	40			40			40		
Akron Un. Pass. Dep. 1st 4...	J&J 1	85			85			85		
⋮										

Figure 4: Mergent/Moody's Bond Record. This figure illustrates our data collection process from the Mergent/Moody's Bond Record. Panel A shows a scanned segment from the January 1972 issue. Panel B displays the results of the data entry process, which involves two independent freelancers separately entering each data item, followed by a third freelancer comparing these entries. When no discrepancy is detected, the entry is accepted. In cases of discrepancy, the third freelancer consults the original document to make a final determination, effectively breaking the tie.

process involving three freelancers.

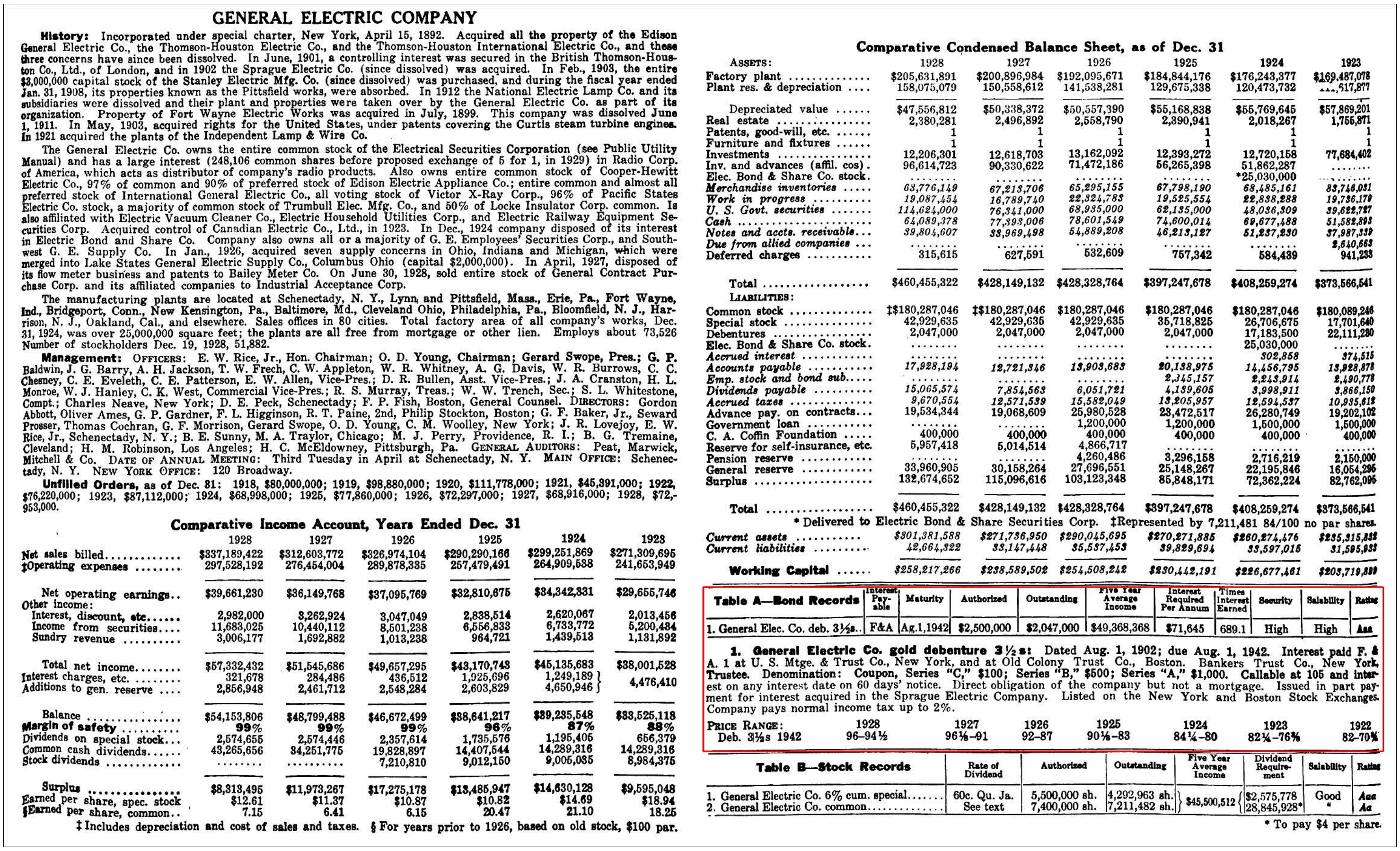
2.2 Bond characteristics from Moody's Manual

In the process of compiling monthly corporate bond quotes, we are able to gather information about the bonds themselves. While some attributes are explicitly given as separate fields (e.g., interest payment dates), most are parsed from bond names (e.g., coupon rate, maturity date). Relying solely on bond names to extract these details, however, is risky and often unreliable. The CFC, S&P Bond Guide, and Mergent/Moody's Bond Record are all print publications, and they frequently adjust how they present bond names by adding or dropping certain terms due to space constraints. Moreover, some bond attributes are not even provided depending on the source. For example, the CFC does not include amounts outstanding or credit ratings. Considering these aspects, it is important to verify the bond information collected from monthly records and to supplement the data with additional characteristics when necessary.

To address this issue, we exploit Moody's Manual, a reference series that contains detailed information on companies and their securities. First published in 1900 and issued annually, Moody's Manual covers various sectors, such as railroads, public utilities, and industrials. For each firm, the manual provides extensive information, including not just company history and business description but also financial statements and details of the securities issued. This enables us to obtain comprehensive bond characteristics by consulting the manual.

It is worth noting that Mergent FISD (Fixed Income Securities Database) is widely used in the literature for similar purposes. While Mergent FISD provides good coverage for bonds that were actively traded in the mid-1990s onward, it offers limited coverage for older bonds. Since our goal is to extend the pricing data on corporate bonds further back over a century, Mergent FISD alone is insufficient. In fact, Mergent FISD has virtually no overlap with our corporate bond records from the CFC and S&P Bond Guide (1895-1972). A large portion of bond records from Mergent/Moody's Bond Record also remains uncovered (1972-2003). We can bridge this gap by referencing Moody's Manual.

Specifically, we collect physical copies of Moody's Manual and digitize them through scanning. Each year, the manual covers a large of number of firms, totaling several thousand pages per year. To illustrate our mapping process, consider the bond record "Gen'l Elec 3 1/2s 1942 opt" observed in Panel A of Figure 2. The process begins by identifying the name of the bond issuer from the bond name. Here, it is apparent that the issuer is General Electric (GE). For each year we observe this bond, we search for GE from that year's Moody's Manual. Figure 5 displays the relevant section of the 1929 manual for GE (pages 2977-2978). Along with the company's condensed income statement and balance sheet, the "Bond Records" table provides detailed information about its bonds. We identify the bond in question and validate existing



GENERAL ELECTRIC COMPANY

History: Incorporated under special charter, New York, April 15, 1892. Acquired all the property of the Edison General Electric Co., the Thomson-Houston Electric Co., and the Thomson-Houston International Electric Co., and these three concerns have since been dissolved. In June, 1901, a controlling interest was secured in the British Thomson-Houston Co., Ltd., of London, and in 1902 the Sprague Electric Co. (since dissolved) was acquired. In Feb., 1903, the entire $3,000,000 capital stock of the Stanley Electric Mfg. Co. (since dissolved) was purchased, and during the fiscal year ended Jan. 31, 1908, its properties known as the Pittsfield works, were absorbed. In 1912 the National Electric Lamp Co. and its subsidiaries were dissolved and their plant and properties were taken over by the General Electric Co. as part of its organization. Property of Fort Wayne Electric Works was acquired in July, 1899. This company was dissolved June 1, 1911. In May, 1903, acquired rights for the United States, under patents covering the Curtis steam turbine engines. In 1921 acquired the plants of the Independent Lamp & Wire Co.

The General Electric Co. owns the entire common stock of the Electrical Securities Corporation (see Public Utility Manual) and has a large interest (248,106 common shares before proposed exchange of 5 for 1, in 1929) in Radio Corp. of America, which acts as distributor of company's radio products. Also owns entire common stock of Cooper-Hewitt Electric Co., 97% of common and 90% of preferred stock of Edison Electric Appliance Co.; entire common and almost all preferred stock of International General Electric Co., all voting stock of Victor X-Ray Corp., 96% of Pacific States Electric Co. stock, a majority of common stock of Trumbull Elec. Mfg. Co., and 50% of Locke Insulator Corp. common. Is also affiliated with Electric Vacuum Cleaner Co., Electric Household Utilities Corp., and Electric Railway Equipment Securities Corp. Acquired control of Canadian Electric Co., Ltd., in 1923. In Dec., 1924 company disposed of its interest in Electric Bond and Share Co. Company also owns all or a majority of G. E. Employees' Securities Corp., and Southwest G. E. Supply Co. In Jan., 1926, acquired seven supply concerns in Ohio, Indiana and Michigan, which were merged into Lake States General Electric Supply Co., Columbus Ohio (capital $2,000,000). In April, 1927, disposed of its flow meter business and patents to Bailey Meter Co. On June 30, 1928, sold entire stock of General Contract Purchase Corp. and its affiliated companies to Industrial Acceptance Corp.

The manufacturing plants are located at Schenectady, N. Y., Lynn and Pittsfield, Mass., Erie, Pa., Fort Wayne, Ind., Bridgeport, Conn., New Kensington, Pa., Baltimore, Md., Cleveland Ohio, Philadelphia, Pa., Bloomfield, N. J., Harrison, N. J., Oakland, Cal., and elsewhere. Sales offices in 80 cities. Total factory area of all company's works, Dec. 31, 1924, was over 25,000,000 square feet; the plants are all free from mortgage or other lien. Employs about 73,526 Number of stockholders Dec. 19, 1928, 51,882.

Management: OFFICERS: E. W. Rice, Jr., Hon. Chairman; O. D. Young, Chairman; Gerard Swope, Pres.; G. P. Baldwin, J. G. Barry, A. H. Jackson, T. W. Frech, C. W. Appleton, W. R. Whitney, A. G. Davis, W. R. Burrows, C. C. Chesney, C. E. Eveleth, C. E. Patterson, E. W. Allen, Vice-Pres.; D. R. Bullen, Asst. Vice-Pres.; J. A. Cranston, H. L. Monroe, W. J. Hanley, C. K. West, Commercial Vice-Pres.; R. S. Murray, Treas.; W. W. Trench, Sec.; S. L. Whitestone, Compt.; Charles Neave, New York; D. E. Peck, Schenectady; F. P. Fish, Boston, General Counsel. DIRECTORS: Gordon Abbott, Oliver Ames, G. P. Gardner, F. L. Higginson, R. T. Paine, 2nd, Philip Stockton, Boston; G. F. Baker, Jr., Seward Prosser, Thomas Cochran, G. F. Morrison, Gerard Swope, O. D. Young, C. M. Woolley, New York; J. R. Lovejoy, E. W. Rice, Jr., Schenectady, N. Y.; B. E. Sunny, M. A. Traylor, Chicago; M. J. Perry, Providence, R. I.; B. G. Tremaine, Cleveland; H. M. Robinson, Los Angeles; H. C. McEldowney, Pittsburgh, Pa. GENERAL AUDITORS: Peat, Marwick, Mitchell & Co. DATE OF ANNUAL MEETING: Third Tuesday in April at Schenectady, N. Y. MAIN OFFICE: Schenectady, N. Y. NEW YORK OFFICE: 120 Broadway.

Unfilled Orders, as of Dec. 81: 1918, $80,000,000; 1919, $98,880,000; 1920, $111,778,000; 1921, $45,391,000; 1922, $76,220,000; 1923, $87,112,000; 1924, $68,998,000; 1925, $77,860,000; 1926, $72,297,000; 1927, $68,916,000; 1928, $72,953,000.

Comparative Income Account, Years Ended Dec. 31

	1928	1927	1926	1925	1924	1923
Net sales billed	$337,189,422	$312,603,772	$326,974,104	$290,290,166	$299,251,869	$271,309,695
‡Operating expenses	297,528,192	276,454,004	289,878,335	257,479,491	264,909,538	241,653,949
Net operating earnings	$39,661,230	$36,149,768	$37,095,769	$32,810,675	$34,342,331	$29,655,746
Other income:						
Interest, discount, etc.	2,982,000	3,262,924	3,047,049	2,838,514	2,620,067	2,013,456
Income from securities	11,683,025	10,440,112	8,501,238	6,556,833	6,733,772	5,200,484
Sundry revenue	3,006,177	1,692,882	1,013,238	964,721	1,439,513	1,131,892
Total net income	$57,332,432	$51,545,686	$49,657,295	$43,170,743	$45,135,683	$38,001,528
Interest charges, etc.	321,678	284,486	436,512	1,925,696	1,249,189	4,476,410
Additions to gen. reserve	2,856,948	2,461,712	2,548,284	2,603,829	4,650,946	
Balance	$54,153,806	$48,799,488	$46,672,499	$38,641,217	$39,235,548	$33,525,118
Margin of safety	99%	99%	99%	96%	87%	88%
Dividends on special stock	2,574,655	2,574,446	2,357,614	1,735,576	1,195,406	656,379
Common cash dividends	43,265,656	34,251,775	19,828,897	14,407,544	14,289,316	14,289,316
Stock dividends	....	....	7,210,810	9,012,150	9,005,035	8,984,375
Surplus	$8,313,495	$11,973,267	$17,275,178	$13,485,947	$14,630,128	$9,595,048
Earned per share, spec. stock	$12.61	$11.37	$10.87	$10.82	$14.69	$18.94
§Earned per share, common	7.15	6.41	6.15	20.47	21.10	18.25

‡Includes depreciation and cost of sales and taxes. §For years prior to 1926, based on old stock, $100 par.

Comparative Condensed Balance Sheet, as of Dec. 31

ASSETS:	1928	1927	1926	1925	1924	1923
Factory plant	$205,631,891	$200,896,984	$192,095,671	$184,844,176	$176,243,377	$169,487,078
Plant res. & depreciation	158,075,079	150,558,612	141,538,281	129,675,338	120,473,732	[illegible],517,877
Depreciated value	$47,556,812	$50,338,372	$50,557,390	$55,168,838	$55,769,645	$57,869,201
Real estate	2,380,281	2,496,892	2,558,790	2,390,941	2,018,267	1,755,871
Patents, good-will, etc.	1	1	1	1	1	1
Furniture and fixtures	1	1	1	1	1	1
Investments	12,206,301	12,618,703	13,162,092	12,393,272	12,720,158	77,684,402
Inv. and advances (affil. cos)	96,614,723	90,330,622	71,472,186	56,265,398	51,862,287	....
Elec. Bond & Share Co. stock	....	....	....	....	*25,030,000	....
Merchandise inventories	*63,776,149*	*67,213,706*	*65,295,155*	*67,798,190*	*68,485,161*	*83,746,031*
Work in progress	*19,087,454*	*16,789,740*	*22,324,783*	*19,525,554*	*22,838,288*	*19,736,179*
U. S. Govt. securities	*114,624,000*	*76,341,000*	*68,935,000*	*62,135,000*	*48,036,309*	*39,622,727*
Cash	*64,089,378*	*77,393,006*	*78,601,549*	*74,600,014*	*69,677,488*	*51,582,893*
Notes and accts. receivable	*39,804,607*	*33,969,498*	*54,889,208*	*46,213,127*	*51,237,230*	*37,987,339*
Due from allied companies	....	....	....	....	....	*2,640,663*
Deferred charges	315,615	627,591	532,609	757,342	584,439	941,233
Total	$460,455,322	$428,149,132	$428,328,764	$397,247,678	$408,259,274	$373,566,541
LIABILITIES:						
Common stock	‡$180,287,046	‡$180,287,046	$180,287,046	$180,287,046	$180,287,046	$180,089,246
Special stock	42,929,635	42,929,635	42,929,635	35,718,825	26,706,675	17,701,640
Debentures	2,047,000	2,047,000	2,047,000	2,047,000	17,183,500	22,111,230
Elec. Bond & Share Co. stock	....	....	....	....	25,030,000	....
Accrued interest	....	....	....	....	*302,858*	*374,515*
Accounts payable	*17,928,194*	*12,721,346*	*13,903,683*	*20,138,975*	*14,456,795*	*13,928,878*
Emp. stock and bond sub.	....	....	....	*2,345,157*	*2,243,914*	*2,490,778*
Dividends payable	*15,065,574*	*7,854,563*	*6,051,721*	*4,139,605*	*3,998,911*	*3,866,150*
Accrued taxes	*9,670,554*	*12,571,539*	*15,582,049*	*13,205,957*	*12,594,537*	*10,935,612*
Advance pay. on contracts	19,534,344	19,068,609	25,980,528	23,472,517	26,280,749	19,202,102
Government loan	....	....	1,200,000	1,200,000	1,500,000	1,500,000
C. A. Coffin Foundation	400,000	400,000	400,000	400,000	400,000	400,000
Reserve for self-insurance, etc.	5,957,418	5,014,514	4,866,717	....	....	....
Pension reserve	....	....	4,260,486	3,296,158	2,716,219	2,150,000
General reserve	33,960,905	30,158,264	27,696,551	25,148,267	22,195,846	16,054,295
Surplus	132,674,652	115,096,616	103,123,348	85,848,171	72,362,224	82,762,095
Total	$460,455,322	$428,149,132	$428,328,764	$397,247,678	$408,259,274	$373,566,541

* Delivered to Electric Bond & Share Securities Corp. ‡Represented by 7,211,481 84/100 no par shares.

	1928	1927	1926	1925	1924	1923
Current assets	*$301,381,588*	*$271,736,950*	*$290,045,695*	*$270,271,885*	*$260,274,476*	*$235,315,832*
Current liabilities	*42,664,322*	*33,147,448*	*35,537,453*	*39,829,694*	*33,597,015*	*31,595,933*
Working Capital	*$258,217,266*	*$238,589,502*	*$254,508,242*	*$230,442,191*	*$226,677,461*	*$203,719,899*

Table A—Bond Records	Interest Payable	Maturity	Authorized	Outstanding	Five Year Average Income	Interest Required Per Annum	Times Interest Earned	Security	Salability	Rating
1. General Elec. Co. deb. 3½s	F&A	Ag.1,1942	$2,500,000	$2,047,000	$49,368,368	$71,645	689.1	High	High	Aaa

1. General Electric Co. gold debenture 3½s: Dated Aug. 1, 1902; due Aug. 1, 1942. Interest paid F. & A. 1 at U. S. Mtge. & Trust Co., New York, and at Old Colony Trust Co., Boston. Bankers Trust Co., New York, Trustee. Denomination: Coupon, Series "C," $100; Series "B," $500; Series "A," $1,000. Callable at 105 and interest on any interest date on 60 days' notice. Direct obligation of the company but not a mortgage. Issued in part payment for interest acquired in the Sprague Electric Company. Listed on the New York and Boston Stock Exchanges. Company pays normal income tax up to 2%.

PRICE RANGE:	1928	1927	1926	1925	1924	1923	1922
Deb. 3½s 1942	96–94½	96⅛–91	92–87	90⅛–83	84¼–80	82¼–76⅝	82–70⅝

Table B—Stock Records	Rate of Dividend	Authorized	Outstanding	Five Year Average Income	Dividend Requirement	Salability	Rating
1. General Electric Co. 6% cum. special	60c. Qu. Ja.	5,500,000 sh.	4,292,963 sh.	$45,500,512	$2,575,778	Good	Aaa
2. General Electric Co. common	See text	7,400,000 sh.	7,211,482 sh.		28,845,928*	"	Aa

* To pay $4 per share.

Figure 5: Moody's Manual – General Electric. This figure displays a section of the 1929 Moody's Manual containing information about General Electric. The excerpt provides details on the company's history, business description, condensed financial statements, and issued securities. Specifically, bond details are listed in the "Bond Records" table and its caption. From this, we collect and verify key bond attributes, including the issue date, dated date, maturity date, coupon details (rate, type, payment dates), denomination currency, issuer domicile, collateral status, embedded options, amount outstanding, and Moody's credit rating.

data. Additionally, we collect new bond attributes, if not already present, such as issue date, dated date, maturity date, coupon details (rate, type, payment dates), denomination currency, issuer domicile, collateral status, embedded options, amount outstanding, and Moody's credit rating. Note that the last two items can vary over time, so we update them annually. A couple of remarks are in order. (i) The amount outstanding of a given bond can also be obtained from the S&P Bond Guide and Mergent/Moody's Bond Record (the CFC does not report this). Occasionally, the amount outstanding reported in Moody's Manual differs from the values in these other sources. These discrepancies are typically due to one of the data sources lagging a few months in revising the amount. Historically, many bonds were issued with sinking fund provisions, leading to a gradual decrease in the amount outstanding. To capture this partial debt retirement in a timely fashion, we choose the smaller value of the two. (ii) Although the S&P Bond Guide and Mergent/Moody's Bond Record provide monthly credit ratings, we currently do not collect them and instead rely on the annual crediting rating information from Moody's Manual. This is in place to maintain consistency across the three data sources, given that the CFC does not provide any credit ratings and that the S&P Bond Guide and Mergent/Moody's Bond Record use different credit rating systems.

Moody's Manual proves valuable for enriching monthly corporate bond records, not just from the three new print sources but also from existing datasets. In the case of WRDS, the sample is relatively recent and is fully covered by Mergent FISD. However, a substantial portion of observations in the Lehman Brothers dataset is still not matched with Mergent FISD. For these unmatched bonds, we rely on Moody's Manual to manually supplement their characteristics.

2.3 Bond filtering and time series aggregation

We construct our historical corporate bond database by consolidating monthly bond records from five different sources. As an initial step, we apply conventional filters to the monthly corporate bond records, following Dickerson, Mueller, and Robotti (2023). Specifically, we only retain corporate bonds that are issued by domestic entities domiciled in the U.S. and are denominated in U.S. dollars. We also exclude exotic instruments such as structured notes, equity-linked notes, mortgage-backed securities, and asset-backed securities.[10] Unlike Dickerson, Mueller, and Robotti (2023), we keep bonds with variable coupons or nonstandard coupon frequencies as well as convertible bonds at this stage, as they are legitimate corporate bonds and should remain in the database. However, consistent with the conventions in the literature, we drop them from our empirical analysis in Sections 3 and 4.

Next, we generate and assign a unique ID number to each bond. We create our own ID

[10]Dickerson, Mueller, and Robotti (2023) additionally discard bonds that are privately placed or traded under Rule 144A, based on the relevant fields provided by Mergent FISD. We apply these filters where possible: all bonds in WRDS and some in Lehman Brothers and Mergent/Moody's Bond Record are matched with Mergent FISD. For other bonds, we cannot determine whether they are privately placed/traded. However, such bonds are unlikely to have been reported in newspapers like the CFC or by credit rating agencies such as S&P and Moody's in the first place.

system due to the lack of consistent bond identifiers within and across these sources. Neither the CFC nor the S&P Bond Guide provides any bond identifiers. Although CUSIP (Committee on Uniform Security Identification Procedures) identifiers were introduced in 1964, they are not reported in Mergent/Moody's Bond Record until 1991.

Despite this, we are able to track bonds over time and construct their time series based on their names. This process is not always straightforward. We see changes in bond names when the issuers change their names or go through mergers and acquisitions. Bond names can also be printed differently because of variations in abbreviations, acronyms, or the omission of certain terms. To ensure data integrity, we review these cases individually. Specifically, when a bond name suddenly disappears in one month, we examine the list of newly appeared bonds for that month to check if one of them is actually the same bond, just under a different name. This verification is based on comparing the characteristics of the original and candidate bonds, with the help of Moody's Manual. When building a time series for the CFC, S&P Bond Guide, and Mergent/Moody's Bond Record, it is important to keep in mind that the publication month does not correspond to the data month. Bond records from these print sources reflect values at the end of the preceding month. For example, the March 1928 issue of the CFC contains bond quotes observed at the end of February 1928. Therefore, we need to shift the publication month backward by one month to align with the actual data month.

Source	Sample period	# Unique bonds	# Observations	# Bonds per month
CFC	04/1895 - 11/1944	13,075	1,624,719	2,726
S&P Bond Guide	05/1943 - 11/1972	10,472	943,437	2,635
Mergent/Moody's Bond Record	12/1971 - 11/2003	26,743	2,479,034	6,456
Lehman Brothers	01/1973 - 03/1998	17,146	1,686,978	5,568
WRDS	07/2002 - 06/2022	70,758	2,719,496	11,331
Aggregate database	04/1895 - 06/2022	109,570	7,739,314	5,068

Table 1: Descriptive characteristics by source. The table presents descriptive characteristics of our corporate bond database, which spans from April 1895 to June 2022 and combines data from five major sources: CFC, S&P Bond Guide, Mergent/Moody's Bond Record, Lehman Brothers, and WRDS. For each source, it shows the sample period, total number of unique bonds, total number of observations, and the average number of bonds per month. The final row represents the aggregated corporate bond database.

Table 1 provides descriptive statistics for the aggregate corporate bond database, which spans over a century, from April 1895 to June 2022. The three new hand-collected sources significantly improve the breadth and depth of the database. The CFC and S&P Bond Guide extend the historical coverage of our database, providing observations from the early 1900s and even the late 1800s. Although Mergent/Moody's Bond Record largely overlaps with the Lehman Brothers dataset in terms of sample period, it is extremely useful for addressing the data gap between Lehman Brothers and WRDS from 1998 to 2002.

The resulting corporate bond database includes nearly 110,000 unique bonds and 8 million total observations. On average, the number of bonds per month varies by source. The number of bonds per month is around 2,500 for the CFC and S&P Bond Guide, and we see this number rise to over 5,000 for Mergent/Moody's and Lehman Brothers datasets. WRDS, which covers the most recent period, records the highest average at about 8,500 bonds per month.

2.4 Merging with the CRSP stock database

Providing stock prices for all publicly traded firms in the U.S. since 1926, the CRSP stock database is widely regarded as the gold standard for stock market data and related research. We claim that the corporate bond database we construct can serve as a "CRSP" for corporate bonds, offering a large cross-section and an extensive time series of corporate bond prices, with coverage comparable to that of CRSP. In this section, we describe the process of merging the two databases. Since not all public firms issue bonds and not all bonds are issued by public firms, the intersection of the two databases is smaller than the two originals. Nevertheless, we obtain a significant number of matched firms and bonds. Through this merged database, we are able to perform a joint analysis of stocks and bonds issued by the same firms and gain a more complete understanding of how corporate securities are priced and interact with one another.

For each year, we first identify firms whose stocks are publicly traded. In the CRSP database, firms are uniquely identified by their PERMCOs. We consult Moody's Manual and make a complete list of bonds issued by these firms and their subsidiaries. The bonds identified through Moody's Manual are then manually compared with records in our corporate bond database. If a match is found, the corresponding bond ID is assigned to the firm's PERMCO.

This approach ensures that each PERMCO can be matched with multiple bonds if the firm had issued more than one bond. We flag an observation if a parent company's PERMCO is mapped into bonds issued by its subsidiaries. We repeat this every year from 1926 up to 2001, as the WRDS portion of our database (starting in 2002) already comes merged with CRSP.

Table 2 presents the summary statistics of the merged database, averaged over five-year intervals. It reports the average number of CRSP-listed firms per year and the subset whose bonds are matched in our corporate bond database. Additionally, it also provides the average number of matched firms after excluding bonds issued by subsidiaries, distinguishing between parent company and subsidiary-level activities. In the early part of the sample, approximately 40% of CRSP-listed firms have bonds captured by our database (e.g., 302 out of 701 firms in 1926-1930). Although the proportion of matched firms declines to around 10% toward the end of the sample period (e.g., 706 out of 7,496 firms in 2001-2005), the absolute numbers of CRSP firms and matched firms both grow substantially over time.

	# CRSP firms per year			# Bonds per year			# Bonds per firm	
Sample period	All	Merged	No subs	All	Merged	No subs	Merged	No subs
1926 - 1930	701	302	273	3,685	1,964	891	6.50	3.26
1931 - 1935	794	306	281	3,838	1,945	873	6.35	3.10
1936 - 1940	855	265	239	3,451	1,546	723	5.83	3.03
1941 - 1945	885	225	196	3,109	1,189	570	5.29	2.91
1946 - 1950	1,018	254	220	2,272	1,118	621	4.40	2.82
1951 - 1955	1,108	305	267	2,329	1,430	918	4.68	3.44
1956 - 1960	1,170	417	378	2,559	1,888	1,281	4.52	3.39
1961 - 1965	1,804	540	494	2,855	2,337	1,632	4.33	3.30
1966 - 1970	2,322	657	619	3,512	2,862	2,042	4.36	3.30
1971 - 1975	4,460	908	874	5,114	4,228	3,006	4.66	3.44
1976 - 1980	5,456	875	835	6,211	4,816	3,450	5.50	4.13
1981 - 1985	6,267	860	822	7,098	5,004	3,521	5.82	4.28
1986 - 1990	7,323	830	783	8,418	4,890	3,334	5.89	4.26
1991 - 1995	7,535	774	723	8,810	4,195	2,944	5.42	4.07
1996 - 2000	9,022	818	770	8,386	3,730	2,826	4.56	3.67
2001 - 2005	7,622	763	736	10,218	3,903	3,002	5.12	4.08

Table 2: Summary statistics of matched CRSP firms and bonds. This table presents summary statistics of the merged database, averaged over five-year intervals. The "# CRSP firms per year" columns report the average number of CRSP-listed firms per year ("Total"), the subset matched with the corporate bond database ("Merged"), and the subset excluding bonds issued by subsidiaries ("No subs"). The "# Bonds per year" columns show the average number of bonds per year from the corporate bond database ("Total"), the subset matched with CRSP ("Merged"), and the subset excluding bonds issued by subsidiaries ("No subs"). The "# Bonds per firm" columns display the average number of bonds per matched firm, both including ("Merged") and excluding subsidiaries ("No subs").

The table also reports the average number of bonds per year in our corporate bond database and the subset matched with CRSP. Similarly, it provides the average excluding bonds issued by subsidiaries. The results show that a significant portion of bonds in our database are issued by public firms in CRSP and their subsidiaries, consistently exceeding 38% in all intervals. Initially, around 50% of bonds are matched, with this rate steadily rising beyond 80% in the mid-sample period (e.g., 82% in 1961-1965), before declining in the 1990s and reverting to around 40% by the 2000s. Lastly, the table shows the average number of bonds per matched firm, both including and excluding subsidiaries. This ratio remains relatively stable across the sample period, ranging from 4 to 6 bonds per firm (or 2 to 4 bonds, excluding subsidiaries).

3 Corporate bond yields and credit spreads

In this section, we investigate the historical time series of corporate bond yields and credit spreads, starting in 1895. Before converting bond prices into yields, we apply several additional filters standard in the literature. We exclude bonds with convertibility, variable coupons, and nonstandard coupon frequencies – features that complicate conventional yield calculation – and focus on fixed-coupon (including zero-coupon) debentures with no specific collateral attached. We also drop bonds issued by financials.

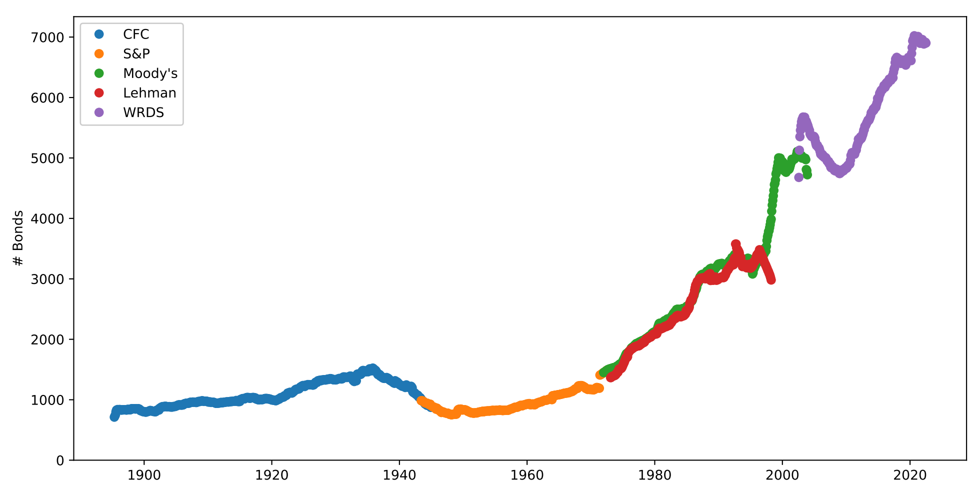


Figure 6: Monthly bond counts by data source. This figure presents the monthly time series of bond counts, with each data source represented by a distinct color: CFC (blue), S&P Bond Guide (orange), Mergent/Moody's Bond Record (green), Lehman Brothers (red), and WRDS (purple). The sample period is from May 1895 to June 2022.

Figure 6 presents the monthly time series of final bond counts by data source. Despite patching together data from different sources, the figure shows a smooth curve, free of extreme dips or spikes, suggesting no systematic loss or gain of bonds when we move from one dataset to another. The near-perfect overlap between the time series of Mergent/Moody's Bond Record and Lehman Brothers further validates the consistency of our data collection. Together, these patterns provide a reassuring sanity check before proceeding with our analysis.

3.1 Corporate bond yields in historical perspective

Figure 7 shows the median yield to maturity for bonds in each month over our sample period from May 1895 to June 2022, providing a long-term perspective on the evolution of bond yields. Similar to Figure 6, the data from each source are represented by a distinct color. The overlapping periods between these sources offer an opportunity to evaluate consistency across different datasets. Notably, during the considerable overlap between the Lehman Brothers dataset and Mergent/Moody's Bond Record, the patterns of median yields align closely, suggesting that the data from different providers are consistent in capturing corporate yield trends over time.

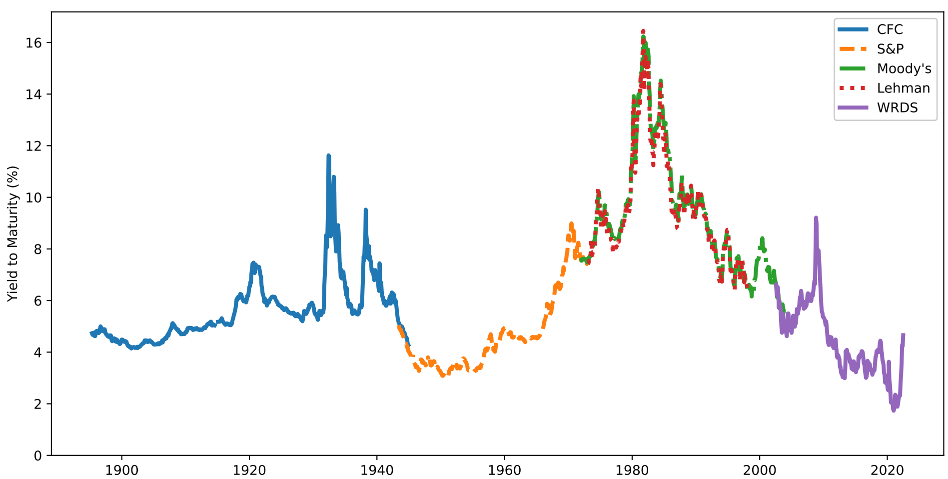


Figure 7: Time series of median bond yields. This figure presents the monthly time series of median bond yields, with each data source represented by a distinct color: CFC (blue), S&P Bond Guide (orange), Mergent/Moody's Bond Record (green), Lehman Brothers (red), and WRDS (purple). All values are expressed in percentages. The sample period is from May 1895 to June 2022.

The figure reveals several important historical patterns in bond yields. A marked increase is observed before and at the beginning of the 1920s, reflecting higher borrowing costs during the post-World War I recessions of 1918-1919 and 1920-1921. This is followed by a sharp spike in yields during the 1930s, coinciding with the economic downturns of the Great Depression. The most pronounced surge starts in the mid-1960s and peaks in 1980, driven by the persistent inflationary pressures during this period. This substantial rise in yields is followed by a gradual decrease over the subsequent decades, reflecting a period of diminishing inflation and prolonged declines in interest rates.

The figure also highlights notable spikes in bond yields associated with recent major economic downturns. For instance, an increase is evident around the 2000 dot-com bubble burst. More significantly, the 2008 global financial crisis led to a substantial rise in yields, corresponding to heightened credit risk in financial markets. More recently, bond yields spiked during the COVID-19 pandemic in 2020 as markets initially reacted to uncertainty, though they quickly stabilized with monetary easing. A sharp increase also occurred in 2022, stemming from the recent inflationary episode and the subsequent tightening of monetary policy by the Federal Reserve.

3.2 A long-term view of credit spreads

Bond yields are influenced by several factors, including the short-term risk-free rate, the term premium, and corporate credit risk. To isolate the role of credit risk, we compute the credit spread for all bonds in our database, starting in December 1925. This is achieved by subtracting the Treasury yield with the same maturity from the corporate bond yield. We then calculate the median credit spread for each month, as shown in Figure 8. Several notable patterns emerge.

First, credit spreads peaked during the Great Depression, surpassing 8% in the 1930s. This dramatic increase reflects heightened default risk, significant financial instability, and the elevated borrowing costs faced by corporations during this period of economic distress – levels unmatched throughout the entire sample period. Following the 1930s, credit spreads steadily narrowed, reflecting improved economic conditions, until the mid-20th century when the post-war recovery brought renewed growth and stability. However, during economic downturns such

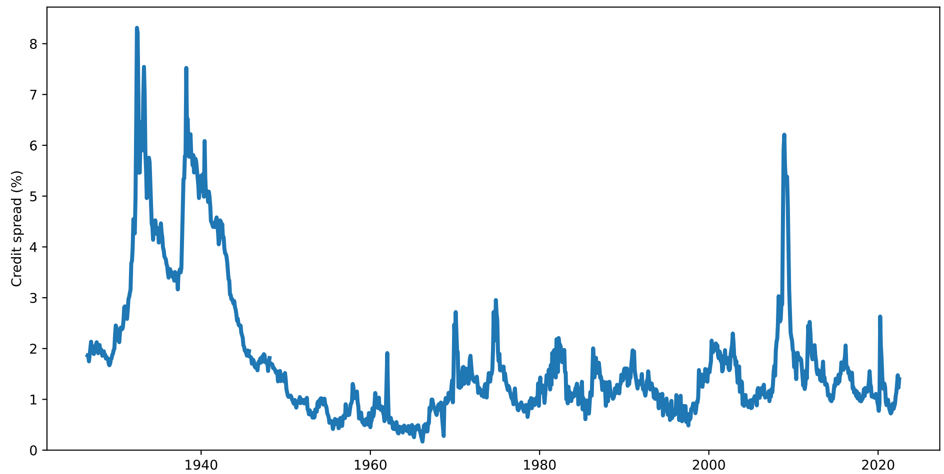


Figure 8: Time series of median credit spreads. This figure presents the monthly time series of median credit spreads. All values are expressed in percentages. The sample period is from July 1926 to June 2022.

as the stagflation of the 1970s and early 1980s, credit spreads experienced marked increases.

The figure also highlights a steep rise in credit spreads during the 2008 global financial crisis, peaking at 6%. This increase reflects tight credit conditions and market turmoil that defined the period, not seen since the Great Depression. Another spike occurred in 2020, coinciding with the onset of the COVID-19 pandemic, though the inflationary period of 2022 resulted in a much more modest increase in credit spreads.

Figure 9 shows the monthly time series of median credit spreads across four rating categories: (i) AAA/AA (solid blue line), (ii) A (dashed orange line), (iii) BBB (dot-dashed green line), and (iv) high-yield (HY) or speculative grade (dotted red line). The figure reveals substantial variation in credit spreads across rating categories, with a clear hierarchy in their levels. Lower-rated bonds, particularly those of speculative grade, consistently exhibit higher credit spreads compared to higher-rated bonds. This disparity becomes particularly pronounced during economic crises, as evidenced during the Great Depression (when high-yield median spreads exceeded 15%), the 2008 global financial crisis, and the COVID-19 pandemic. In contrast, AAA/AA-rated bonds maintain relatively stable and low credit spreads (generally below 2%) even during periods of market turbulence, while A and BBB-rated bonds exhibit moderate volatility that falls between these extremes.

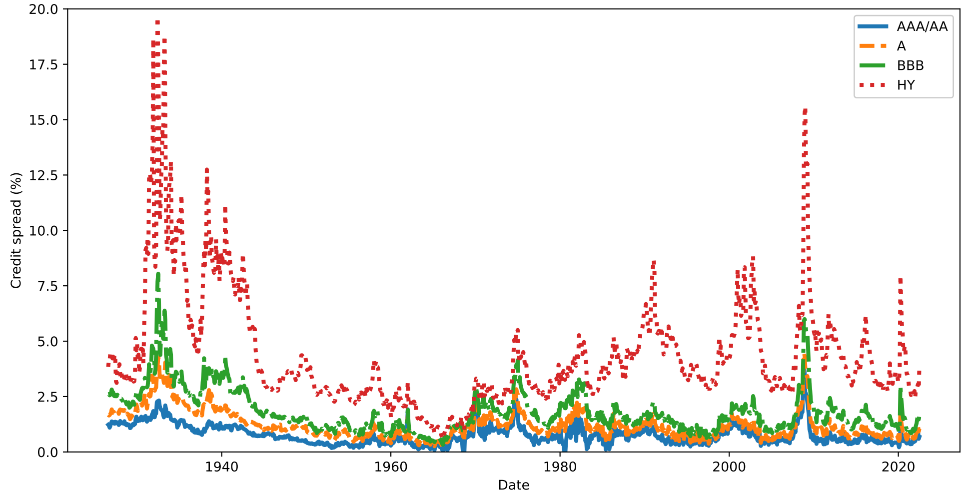


Figure 9: Time series of median credit spreads by credit ratings. This figure presents the monthly time series of median credit spreads across four rating categories: (i) AAA/AA (solid blue line), (ii) A (dashed orange line), (iii) BBB (dot-dashed green line), and (iv) high-yield (HY) or speculative grade (dotted red line). All values are expressed in percentages. The sample period is from July 1926 to June 2022.

3.3 GZ spreads

A seminal paper by Gilchrist and Zakrajšek (2012) proposes a new aggregate corporate credit spread measure, which they call the GZ spread, constructed using a sample of senior unsecured bonds issued by U.S. public non-financial firms. To generate this measure, they first calculate each bond's credit spread as the difference between its yield and the yield of a hypothetical Treasury with identical cash flows. The GZ spread is then calculated as the equally-weighted average of these credit spreads across the entire maturity and credit rating spectrum.

The GZ spread time series has proven particularly valuable for understanding future economic activity. It robustly predicts various macroeconomic indicators, such as industrial production, unemployment, and real GDP growth. Follow-up work by Favara, Gilchrist, Lewis, and Zakrajšek (2016a) shows that the GZ spread also effectively predicts the likelihood of an NBER-dated recession occurring in the next 12 months. The series can be decomposed into a predictable component, reflecting expected defaults, and a residual component called the excess bond premium. Gilchrist and Zakrajšek (2012) demonstrate that this excess bond premium is closely linked to fluctuations in credit supply and the financial system's overall risk-bearing capacity, providing significant information for predicting economic outcomes.

The original series begins in 1973 and has been continuously updated by the Federal Reserve since 2016.[11]

We closely follow the methodology of Gilchrist and Zakrajšek (2012) to extend the time series of the GZ spread using our corporate bond database merged with CRSP. Although our corporate bond database starts in 1895, the GZ spread series cannot go back beyond 1926 (i.e., when CRSP data begin), as it is entirely based on bonds issued by public firms. We exclude bonds issued by financials, bonds with credit spreads below 5 basis points or above 3500 basis points, small corporate issues of less than $1 million, and bonds with time-to-maturities of less than 1 year or more than 30 years.

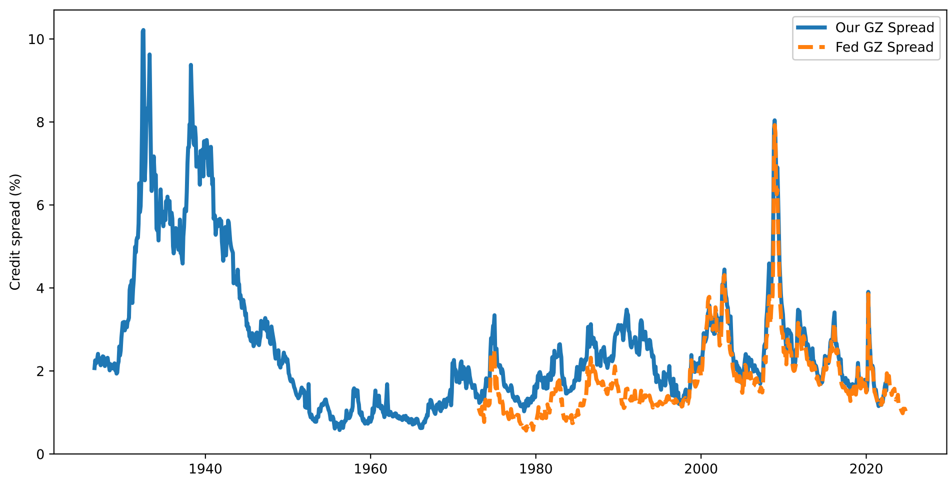


Figure 10: Time series of GZ spreads. This figure presents the monthly time series of the extended GZ spread (solid blue line) alongside the original series from the Federal Reserve for comparison (dashed orange line). All values are expressed in percentages. The sample period is from July 1926 to June 2022.

Figure 10 displays our extended GZ spread series (solid blue line) starting in 1926 and compares it with the original series (dashed orange line) from the Federal Reserve's website. The two series track each other closely post-2000, though some differences appear in the 1980s and 1990s, with an overall correlation of 88%. This gap largely reflects differences in data sources: while Gilchrist and Zakrajšek (2012) construct their series using the Lehman Brothers and Merrill Lynch datasets, our extended series combines print archival sources with the Lehman Brothers data. Indeed, when we construct the GZ spread using only the Lehman

[11] The updated series is available on the Federal Reserve's website. See Favara, Gilchrist, Lewis, and Zakrajšek (2016b) for further details.

Brothers dataset, the correlation between the two series rises to 99%.[12]

Given the strong predictive power of the GZ spread documented in recent samples, we examine whether this predictability extends to our longer historical sample, which encompasses more recessions and extreme business cycle fluctuations, including the Great Depression. Following Favara, Gilchrist, Lewis, and Zakrajšek (2016a), we test the GZ spread's ability to predict future recessions using the probit model:

$$\mathbb{1}\left\{\text{Recession in [t, t+12]}\right\} = \begin{cases} 1 & \text{if } \alpha + \beta_{\text{GZ}} \times \text{GZ}_t + \beta_{\text{TS}} \times \text{TS}_t + \beta_{\text{RFF}} \times \text{RFF}_t + \epsilon_{t+12} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where the dependent variable denotes the NBER recession indicator over the next 12 months. The main independent variable is the GZ spread (GZ_t), with two control variables capturing other financial conditions: the term spread (TS_t), defined as the yield difference between 3-month and 10-year Treasuries, and the real federal funds rate (RFF_t). The error term ϵ_{t+12} is assumed to follow a normal distribution with the distribution function $\Phi(\cdot)$. This model is equivalent to:

$$P\left(\mathbb{1}\left\{\text{Recession in [t, t+12]}\right\} = 1\right) = \Phi\left(\alpha + \beta_{\text{GZ}} \times \text{GZ}_t + \beta_{\text{TS}} \times \text{TS}_t + \beta_{\text{RFF}} \times \text{RFF}_t\right),$$

which directly shows how the GZ spread affects the likelihood of a future recession. We estimate the model via maximum likelihood and report the marginal effects of the three independent variables in Table 3, along with the pseudo R^2. Shown in square brackets are t-statistics, computed using Newey-West standard errors with 12 lags.

In the table, we consider four regression specifications. Specifications (i) and (ii) examine the common sample period (i.e., post-1973) where the original GZ spread and our GZ spread coexist. Both series significantly predict the likelihood of future recessions. Specifically, a 1% (100 basis point) increase in our GZ spread is associated with a 14 percentage point increase in the recession probability. The differences between the two series are minor, though our GZ

[12] We thank Yoshio Nozawa for providing information about the Lehman Brothers dataset and additional guidance for its use.

NBER recession	GZ	TS	RFF	Pseudo R^2
(i) Fed GZ: 1973-2022	0.14	0.12	0.04	0.35
	[3.89]	[4.00]	[2.44]	
(ii) Our GZ: 1973-2022	0.18	0.14	0.03	0.42
	[5.67]	[4.67]	[2.02]	
(iii) Our GZ: 1947-2022	0.15	0.19	0.01	0.23
	[4.32]	[6.13]	[1.05]	
(iv) Our GZ: 1926-2022	0.08	0.17	0.02	0.19
	[3.65]	[5.84]	[2.30]	

Table 3: Predictive power of the GZ spread for future recessions. This table reports the estimation results for the following probit model:

$$\mathbb{1}\,\{\text{Recession in [t, t+12]}\} = \begin{cases} 1 & \text{if } \alpha + \beta_{\text{GZ}} \times \text{GZ}_t + \beta_{\text{TS}} \times \text{TS}_t + \beta_{\text{RFF}} \times \text{RFF}_t + \epsilon_{t+12} > 0, \\ 0 & \text{otherwise,} \end{cases}$$

where the dependent variable denotes the NBER recession indicator over the next 12 months. The main independent variable is the GZ spread (GZ_t), with two control variables capturing other financial conditions: the term spread (TS_t), defined as the yield difference between 3-month and 10-year Treasuries, and the real federal funds rate (RFF_t). The error term ϵ_{t+12} is assumed to follow a normal distribution. We estimate the model via maximum likelihood and report the marginal effects of the three independent variables, along with the pseudo R^2. All t-statistics, shown in square brackets, are computed using Newey-West standard errors with 12 lags.

spread shows slightly stronger predictive power with a higher t-statistic (5.67 vs. 3.89) and a larger pseudo R^2 (0.42 vs. 0.35). Most importantly, specifications (iii) and (iv) demonstrate that the predictive power of our GZ spread remains strong in the post-war sample starting in 1947 as well as when we extend the sample as far as the GZ spread allows.

Taking a step further, we examine the GZ spread's predictive power for future economic activity over our longer historical sample. As in Gilchrist and Zakrajšek (2012), we consider the regression below:

$$y_{t+h} = \alpha + \sum_{k=1}^{p} \beta_k \times y_{t-k} + \beta_{\text{GZ}} \times \text{GZ}_t + \beta_{\text{TS}} \times \text{TS}_t + \beta_{\text{RFF}} \times \text{RFF}_t + \epsilon_{t+h},$$

where y_{t+h} represents the annualized percentage growth rate of two economic variables: industrial production (Panel A) and nonfarm payroll employment (Panel B). For forecast horizon h, the dependent variable y_{t+h} is measured over the $h+1$ periods between times $t-1$ and $t+h$. This timing follows Gilchrist and Zakrajšek (2012), accounting for typical reporting lags that make the growth rate between times $t-1$ and t unobservable at time t. The number of autoregressive terms p is determined by the Akaike Information Criterion. Table 4 reports the estimated regression coefficients for the three independent variables over 3-month and

12-month horizons, together with adjusted R^2 values. The t-statistics in square brackets are calculated using Newey-West standard errors with 12 lags.

	3-month forecasts				12-month forecasts			
Panel A: Industrial production	GZ	TS	RFF	Adj. R^2	GZ	TS	RFF	Adj. R^2
(i) Fed GZ: 1973-2022	-0.41	-0.24	-0.07	0.20	-0.40	-0.40	-0.10	0.27
	[-4.11]	[-2.47]	[-0.84]		[-3.76]	[-3.98]	[-1.03]	
(ii) Our GZ: 1973-2022	-0.49	-0.34	0.03	0.23	-0.46	-0.50	0.01	0.29
	[-4.64]	[-3.09]	[0.3]		[-3.6]	[-4.65]	[0.07]	
(iii) Our GZ: 1947-2022	-0.40	-0.25	0.04	0.21	-0.52	-0.40	0.04	0.27
	[-5.77]	[-3.75]	[0.78]		[-6.03]	[-4.89]	[0.48]	
(iv) Our GZ: 1926-2022	0.02	-0.05	-0.04	0.16	0.11	-0.11	-0.12	0.09
	[0.19]	[-1.35]	[-0.51]		[0.99]	[-1.65]	[-1.31]	
Panel B: Employment	GZ	TS	RFF	Adj. R^2	GZ	TS	RFF	Adj. R^2
(i) Fed GZ: 1973-2022	-0.42	-0.11	-0.08	0.16	-0.55	-0.28	-0.18	0.33
	[-3.76]	[-1.06]	[-0.85]		[-5.38]	[-2.24]	[-1.67]	
(ii) Our GZ: 1973-2022	-0.46	-0.20	0.01	0.18	-0.59	-0.39	-0.06	0.35
	[-4.44]	[-1.76]	[0.09]		[-5.77]	[-3.24]	[-0.52]	
(iii) Our GZ: 1947-2022	-0.37	-0.18	0.05	0.13	-0.54	-0.36	0.07	0.25
	[-3.51]	[-2.36]	[0.82]		[-6.39]	[-4.25]	[0.77]	
(iv) Our GZ: 1939-2022	-0.05	-0.07	-0.03	0.10	-0.00	-0.18	-0.02	0.11
	[-0.44]	[-1.46]	[-0.46]		[-0.01]	[-2.56]	[-0.21]	

Table 4: Predictive power of the GZ spread for future economic activity. This table reports the estimation results for the following regression:

$$y_{t+h} = \alpha + \sum_{k=1}^{p} \beta_k \times y_{t-k} + \beta_{\mathrm{GZ}} \times \mathrm{GZ}_t + \beta_{\mathrm{TS}} \times \mathrm{TS}_t + \beta_{\mathrm{RFF}} \times \mathrm{RFF}_t + \epsilon_{t+h},$$

where y_{t+h} represents the annualized percentage growth rate of two economic variables: industrial production (Panel A) and nonfarm payroll employment (Panel B). For forecast horizon h, the dependent variable y_{t+h} is measured over the $h+1$ periods between times $t-1$ and $t+h$. The main independent variable is the GZ spread (GZ_t), with two control variables capturing other financial conditions: the term spread (TS_t), defined as the yield difference between 3-month and 10-year Treasuries, and the real federal funds rate (RFF_t). The number of autoregressive terms p is determined by the Akaike Information Criterion. The estimated regression coefficients for the three independent variables over 3-month and 12-month horizons, together with adjusted R^2 values, are reported. All t-statistics, shown in square brackets, are computed using Newey-West standard errors with 12 lags.

We again consider four regression specifications. Comparing the first two specifications in the post-1973 sample, our GZ spread robustly predicts declines in the two macroeconomic indicators at both forecast horizons, with slightly larger t-statistics and adjusted R^2 values over the original GZ spread for all cases. While this predictive power persists when we expand our sample to the entire post-war period, the coefficients become statistically insignificant when we extend the sample back as far as the data allow (1926 for industrial production and 1939 for nonfarm payroll employment). This is not too surprising: the relation between the GZ spread and economic indicators can be masked by a realization of an extremely severe and prolonged economic disaster like the Great Depression of the 1930s. Similarly, we observe that the GZ spread's ability to predict future recessions becomes weaker over the full sample in Table 3, although statistical significance still remains in that case. Overall, our analysis provides out-of-sample validation of the GZ spread's economic content while revealing how its predictive power varies across different historical episodes.

4 Corporate bond returns

Having examined historical patterns in corporate bond yields and credit spreads, we now turn to corporate bond returns. This section begins by detailing our methodology for calculating excess returns on individual bonds. We then construct the corporate bond market factor, which allows us to trace the time-series dynamics of the aggregate corporate bond market from 1926 to 2022. Recent empirical evidence strongly supports a single-factor framework akin to the CAPM, where the bond market factor alone captures systematic risk (Dickerson, Mueller, and Robotti, 2023). We examine whether this key finding holds in our extended sample period.

4.1 Calculation of returns

The bond quotes in our archival print sources are mostly clean prices. To obtain the "true" price of the bond that can be used in a transaction (i.e., the dirty price), we add the accrued interest to the quoted price. Formally, let $P_t^{i,\text{Clean}}$ denote the clean price of bond i at time t. The dirty price of the bond $P_t^{i,\text{Dirty}}$ is calculated as $P_t^{i,\text{Dirty}} = P_t^{i,\text{Clean}} + AI_t^i$, where AI_t^i represents the accrued interest from the previous coupon payment date to time t. The holding period return on bond i between time t and time $t+1$ is then given by

$$r_{t+1}^i = \frac{P_{t+1}^{i,\text{Dirty}} + C_{t+1}^i}{P_t^{i,\text{Dirty}}} - 1,$$

where C_{t+1}^i is any coupon payment made during this period.

The quotes we collect from the three print sources are either bid, ask, or sales prices. While most quotes are bids, other types appear frequently, particularly in the earlier months of the sample. Sales prices are available exclusively for exchange-traded securities. As is standard in fixed income markets, quotes are expressed as a fraction of face value. Whenever possible, we compute and report returns based on bid, ask, and sales prices. In calculating returns, we do not mix different price types: bid returns are derived from bid prices, and ask returns from ask prices. When multiple returns are available within a source based on different price types, we take their average. If the same bond's returns can be calculated from two sources due to overlapping data coverage, we take the average of the within-source averages. Note that this procedure differs from first averaging quotes and then calculating returns based on the average quotes. Finally, the excess return is obtained as $rx^i_{t+1} = r^i_{t+1} - r_{f,t}$, where the data on the 1-month risk-free rate $r_{f,t}$ are from Kenneth French's website.

In the process of constructing the return time series, we impose a minimum time to maturity of one year to mitigate potential biases associated with short-term securities, consistent with Dickerson, Mueller, and Robotti (2023). Following their approach, we also apply price filters to eliminate outliers that can distort our results. Specifically, we remove any observation with prices below \$5 or above \$1,000. Most bonds with low prices are defaulted bonds, which are excluded from our analysis through these filters.

4.2 Bond market factor

We construct the bond market factor as the value-weighted average of excess returns on corporate bonds available in our database at each point in time. Following the literature, value weights are based on the amount outstanding for each bond.[13] Figure 11 presents the resulting time series from 1926 to 2022. During this period, the average excess return on the aggregate corporate bond market is 0.36% per month, corresponding to an annual risk premium of 4.32%. This return is accompanied by an annualized volatility of 7.77%, resulting

[13]For most of our sample period, the amount outstanding is readily available. However, this characteristic is not directly provided for bonds from the CFC. Consequently, we rely on information from annual issues of Moody's Manual, as described in Section 2.2. Currently, our analysis for the CFC period is limited to bonds issued by public companies, as hand-collecting data from Moody's Manual is still ongoing for nonpublic firms. We plan to incorporate the full dataset in the next version of the draft.

in a Sharpe ratio of 0.56 – a level comparable to that of the aggregate stock market.

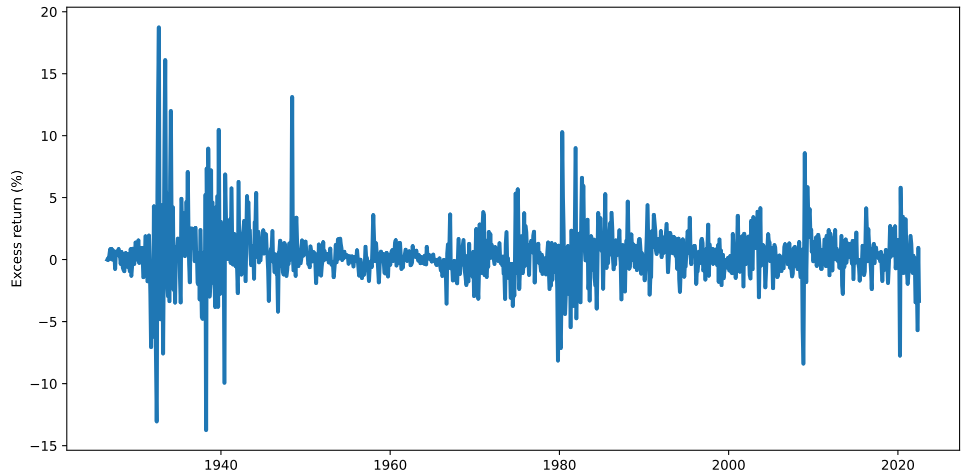


Figure 11: Time series of value-weighted corporate bond excess returns. This figure presents the monthly time series of value-weighted corporate bond excess returns. All values are expressed in percentages. The sample period is from 1926 to 2022.

The figure also demonstrates that our corporate bond database provides an unprecedented historical perspective, making it possible to examine the corporate bond market's response to major economic events. The time series highlights periods of extreme volatility, with pronounced spikes in both positive and negative excess returns during economic crises. The Great Depression of the 1930s stands out, with bond market returns showing extreme fluctuations, reflecting severe financial distress and widespread defaults. Similarly, the stagflationary episode of the 1970s is marked by heightened return volatility, likely driven by inflationary pressures, monetary policy uncertainty, and corporate credit stress, among others. A comparable pattern appears in a more recent sample during the 2008 financial crisis, particularly around the collapse of Lehman Brothers and the subsequent policy interventions.

In contrast to these high-volatility episodes, the 1950s and 1960s exhibit a prolonged period of low return volatility, which coincides with the strong postwar economic expansion. Many of these episodes – including the Great Depression, the aftermath of World War II, the stagflationary period of the 1970s, and the subsequent Volcker disinflation – are not covered by existing corporate bond datasets. The extended time frame of our database allows for a more in-depth assessment of how corporate bond risk and return characteristics evolve across different financial and macroeconomic regimes.

4.3 Is the bond market factor priced?

The bond market factor not only effectively summarizes the aggregate behavior of the corporate bond market in the time series, as shown in Figure 11, but also has been documented as an important priced factor in the cross-section (Dickerson, Mueller, and Robotti, 2023). As the first step toward better understanding the risk-return trade-off in corporate bonds, we examine whether the bond market factor is still priced when extending our sample back to 1926. To this end, we sort corporate bonds into five quintile portfolios based on their exposures to the bond market factor, commonly referred to as pre-ranking betas. These betas are estimated using 36-month rolling windows, with a minimum of 24 months required for an estimate to be included in the analysis. Table 5 presents the results of the bond portfolio sorts. The average excess returns on these portfolios are calculated based on two weighting schemes: equally-weighted (Panel A) and value-weighted (Panel B). The long-short portfolio between Portfolio 5, the most exposed to the market, and Portfolio 1, the least exposed to the market, is also reported and labeled "$5-1$". For each portfolio, we estimate the intercept (α_{CAPMB}) and the slope coefficient (β_{CAPMB}) from a time-series regression, projecting the portfolio return on the bond market factor. The intercept captures abnormal returns relative to the bond market CAPM, with values near zero indicating the model's success in explaining returns. The slope coefficient (β_{CAPMB}), called post-ranking beta, measures the portfolio's exposure to the bond market factor. Note that the subscript "CAPMB" refers to the bond market CAPM. All t-statistics, shown in square brackets, are computed using Newey-West standard errors with 12 lags.

In Panel A, which presents equally-weighted portfolios, the minimum average excess return among these portfolios is 0.30% (Portfolio 2), while the maximum is 0.97% (Portfolio 5). The average excess returns generally follow an increasing trend, with the exception of Portfolio 1, which has an average excess return of 0.48%. The long-short portfolio ($5-1$) shows a statistically significant average return, consistent with predictions from the CAPM. As expected, the

	1	2	3	4	5	$5-1$
Panel A: Equally-weighted portfolios						
Excess return	0.48	0.30	0.35	0.48	0.97	0.49
	[6.12]	[4.33]	[4.39]	[4.28]	[5.39]	[3.24]
α_{CAPMB}	0.33	0.16	0.10	0.04	0.33	0.00
	[5.46]	[2.86]	[2.63]	[1.27]	[3.82]	[-0.04]
β_{CAPMB}	0.40	0.41	0.68	1.21	1.78	1.37
	[9.54]	[7.66]	[14.63]	[18.23]	[21.24]	[14.19]
Panel B: Value-weighted portfolios						
	1	2	3	4	5	$5-1$
Excess return	0.38	0.26	0.23	0.37	0.78	0.39
	[5.22]	[3.89]	[3.29]	[3.41]	[4.43]	[2.54]
α_{CAPMB}	0.23	0.10	-0.02	-0.10	0.07	-0.16
	[4.23]	[2.22]	[-0.70]	[-3.96]	[1.02]	[-1.48]
β_{CAPMB}	0.42	0.43	0.69	1.31	1.94	1.53
	[6.92]	[6.90]	[15.27]	[13.13]	[20.50]	[11.89]

Table 5: Bond portfolio sorts. This table sorts corporate bonds into five quintile portfolios based on their exposures to the bond market factor. These exposures (pre-ranking betas) are estimated using rolling 60-month windows, with a minimum of 36 months required for an estimate to be included in the analysis. Panel A presents the results for equally-weighted portfolios, while Panel B presents the results for value-weighted portfolios. Reported for each portfolio are the average excess return, the intercept (α_{CAPMB}), and the slope coefficient (β_{CAPMB}) from a time-series regression, projecting the portfolio return on the bond market factor. The long-short portfolio between Portfolio 5 (most exposed) and Portfolio 1 (least exposed) is also reported and labeled "$5-1$". All t-statistics, shown in square brackets, are computed using Newey-West standard errors with 12 lags.

betas of these five portfolios increase uniformly. However, some portfolios report statistically significant positive alphas, with Portfolios 1 and 5 having the highest alpha at 0.33%.

Since the results from Panel A might be driven by bonds with small sizes, Panel B repeats the analysis using value-weighted portfolios. The point estimates of average excess returns in the value-weighted portfolios exhibit a similar upward trend, except for Portfolio 1. The long-short portfolio $(5-1)$ demonstrates a positive and statistically significant average excess return, indicating that the bond market factor is priced and that a higher market exposure leads to a higher average return. After controlling for portfolio betas, we find that CAPM alphas follow a decreasing trend, except for Portfolio 5. While the alpha for the long-short portfolio $(5-1)$ is statistically insignificant (-0.16%), Portfolio 1 has a positive and statistically significant alpha (0.23%), whereas Portfolio 4 has a negative and statistically significant alpha (-0.10%).

To visualize this pattern, Figure 12 plots the excess returns of the five value-weighted portfolios against their betas, overlaid with two security market lines. The blue solid line represents the theoretical line that connects the risk-free asset (zero excess return and zero beta) and the bond market portfolio (average market return and unit beta). The orange dashed line represents the empirical counterpart, obtained by fitting a line to the five portfolios. We observe that the security market line is flatter in the data than predicted by the model, a pattern reminiscent of findings from stock portfolio sorts (e.g., Jensen, Black, and Scholes, 1972). Nonetheless, the two lines are much closer for the bond market, suggesting that the CAPM performs relatively better for bonds than stocks.[14]

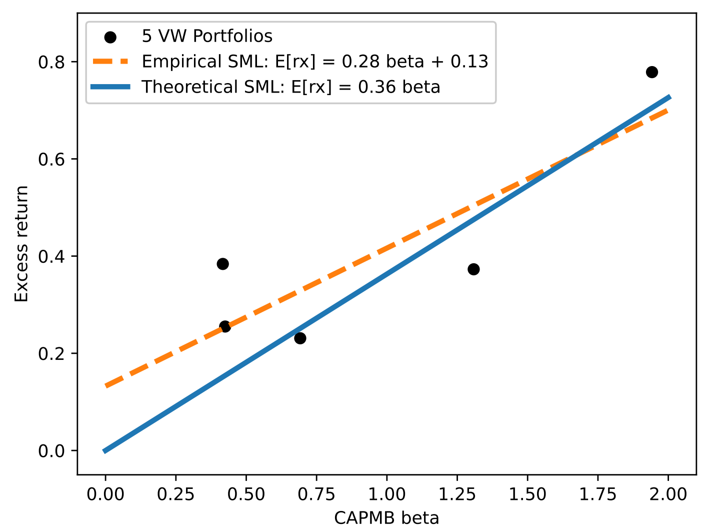


Figure 12: Empirical and theoretical security market lines. This figure plots the excess returns of the five value-weighted portfolios against their betas, overlaid with two security market lines. The black solid line represents the theoretical line that connects the risk-free asset (zero excess return and zero beta) with the bond market portfolio (average market return and unit beta). The red dashed line represents the empirical counterpart, which is the fitted line for the five portfolios.

5 Priced risks in corporate bonds

Section 4 has established that the bond market factor captures an important source of systematic risk in the corporate bond market. This leads us to a key follow-up question: Are there any additional factors that are priced beyond the bond market factor?

While extensive research has uncovered multifactor structures in stock returns that go

[14]For a summary of CAPM performance in the stock market, see Fama and French (2004).

beyond the single-factor CAPM, similar research on corporate bond returns remains in its early stages. Despite the development of various multifactor models, spurred by the well-documented failure of the CAPM in explaining the cross-section of stock returns, these additional factors have shown only limited and mixed success when applied to corporate bond returns. In response, Bai, Bali, and Wen (2019) argue that corporate bond returns also feature a multifactor structure, proposing a four-factor model that includes the (i) bond market factor, (ii) downside risk factor, (iii) credit risk factor, and (iv) liquidity risk factor. This model has since been widely adopted as the benchmark for assessing risk in corporate bonds. However, Dickerson, Mueller, and Robotti (2023) reassess this framework and find that after addressing lead-lag errors and issues with the truncation of the bond market factor, the incremental explanatory power of the other three factors largely vanishes. This result essentially brings the literature back to square one.

What makes Dickerson, Mueller, and Robotti (2023)'s findings even more striking is their strong support for the bond market CAPM. After a comprehensive analysis, they conclude that neither the bond factors previously proposed by Bai, Bali, and Wen (2019) nor prominent stock and macro factors (known to be priced in the cross-section of stock returns) appear to be significantly priced in the cross-section of corporate bond returns. These factors add little explanatory power beyond the bond market factor, leaving the bond market CAPM as the dominant model for the corporate bond market. This conclusion is quite puzzling from any theoretical perspective. If stock/macro factors effectively capture systematic risk to the extent that they are priced in the cross-section of stock returns, why aren't they priced in the cross-section of corporate bond returns? Unless the two markets are significantly segmented or the corporate bond market has substantial frictions masking the true risk-return relation, it is difficult to understand this pattern based on any theory. Can the differences be attributed merely to the fact that stocks and bonds are issued with different seniority and payoff structures? Identifying a common factor structure that bridges both markets remains an even more difficult challenge, which continues to elude us.

We propose a different hypothesis: previous studies in the corporate bond literature rely on a relatively short sample (typically starting in 2002 or later), inevitably limiting the power of statistical tests. For example, Dickerson, Mueller, and Robotti (2023) use a monthly sample from August 2004 to December 2016, with only 149 monthly observations. Given that the pricing data for corporate bonds are potentially noisy, we cannot definitively tell whether the lack of significance indicates that a given factor is unimportant in explaining the cross-section or simply reflects weak statistical power.

Leveraging our new corporate bond database with a much longer sample period, this section revisits these critical empirical issues in asset pricing with greater statistical power. Our analysis yields substantially different conclusions about priced risks in corporate bonds.

5.1 Fama-MacBeth procedure

We implement the Fama and MacBeth (1973) two-pass regression procedure for various linear factor models. Consider a model where K factors $\{f_k \mid k = 1, \cdots, K\}$ capture systematic risk. In the first pass, we estimate bond-level exposures (i.e., pre-ranking betas) by running the following time series regression using past 36-month rolling windows, requiring at least 24 monthly observations:

$$rx_t^i = \beta_0 + \beta_1^i f_{1,t} + \cdots \beta_K^i f_{K,t} + \epsilon_{i,t}. \tag{1}$$

For each factor f_k, we sort bonds into five equally-weighted portfolios based on their pre-ranking betas. Using the full sample (rather than rolling windows), we estimate post-ranking betas for each of these sorted portfolios using the same specification as equation (1). This procedure yields $5 \times K$ post-ranking betas. We then assign these portfolio-level betas to individual bonds based on their portfolio assignment each month. While the post-ranking portfolio betas are constant, individual bonds can receive different beta values over time as they move between portfolios, depending on their pre-ranking beta estimates. This standard approach of using post-ranking portfolio betas at the bond level helps mitigate measurement errors in security-level betas and attenuation biases in estimating risk premia (see Fama and French, 1992; Goldberg and Nozawa, 2021; Dickerson, Mueller, and Robotti, 2023; van Binsbergen, Nozawa, and Schwert, 2024).

Using these bond-level betas, we run the second-pass cross-sectional regression at each month t:

$$rx^i_{t+1} = \gamma_0 + \beta^{i,\text{Post}}_{1,t}\gamma_{1,t} + \cdots \beta^{i,\text{Post}}_{K,t}\gamma_{K,t} + u_{i,t},$$

where the factor risk premium for f_k is estimated by $\gamma_k = \frac{1}{T}\sum_{t=1}^{T}\gamma_{k,t}$. We calculate the Fama-MacBeth t-statistics using Newey-West standard errors with 12 lags.

5.2 Traded bond factors

We begin our analysis by testing well-known bond factors. Table 6 considers three factor models. The first model, "CAMPB," is the bond market CAPM where the bond market factor (MKTB) serves as the sole factor. We call the second model "Modified BBW" as it uses only the first three factors from Bai, Bali, and Wen (2019)'s four-factor model: the bond market factor, downside risk factor (DRF), and credit risk factor (CRF).[15] We omit the fourth factor, the liquidity risk factor (LRF), because its construction relies on an illiquidity measure that requires daily data; the pre-WRDS portion of our return series is available only at a monthly frequency.[16] Lastly, "DEFTERM" refers to a two-factor model with the default return spread (DEF) and the term return spread (TERM), as proposed by Fama and French (1993).[17]

[15]Bai, Bali, and Wen (2019) and Dickerson, Mueller, and Robotti (2023) construct these bond factors by independently sorting bonds into 5 × 5 portfolios based on two characteristics. The DRF is the long-short portfolio return based on the 5% Value at Risk, averaged across different credit rating quintiles. The CRF is the equally-weighted average of three long-short portfolio returns based on credit ratings – CRF_{VaR5} (averaged across five Value-at-Risk portfolios), $\text{CRF}_{\text{ILLIQ}}$ (averaged across five illiquidity portfolios), and CRF_{REV} (averaged across five reversal portfolios). We follow this methodology exactly for the DRF. For the CRF, we simply take the average between CRF_{VaR5} and CRF_{REV} because $\text{CRF}_{\text{ILLIQ}}$ cannot be constructed without daily return data. Nevertheless, we confirm that our CRF time series constructed without $\text{CRF}_{\text{ILLIQ}}$ maintains a very high correlation around 98% with the original series in the post-2002 WRDS sample period.

[16]The omission of the liquidity risk factor is less critical for our analysis, as it already shows marginal significance and explanatory power beyond the bond market factor in Dickerson, Mueller, and Robotti (2023)'s relatively short sample. We focus instead on the downside and credit risk factors, which they find to be insignificant.

[17]We obtain the time series of these two factors from Amit Goyal's website.

	CAPMB		Modified BBW				DEFTERM		
	γ_0	γ_{MKTB}	γ_0	γ_{MKTB}	γ_{DRF}	γ_{CRF}	γ_0	γ_{DEF}	γ_{TERM}
Panel A: Short sample	07/2002-12/2016		07/2002-12/2016				07/2002-12/2016		
Estimate	0.10	0.41	0.20	0.28	0.07	0.22	0.30	0.32	0.13
	[0.58]	[1.18]	[2.32]	[1.3]	[0.37]	[0.71]	[2.73]	[0.64]	[0.29]
Adj. R^2	0.05		0.11				0.09		
Panel B: Long sample	07/1926-06/2022		07/1926-06/2022				07/1926-06/2022		
Estimate	0.17	0.39	0.30	0.24	1.37	0.72	0.10	0.45	0.15
	[2.47]	[3.84]	[3.82]	[2.53]	[4.98]	[4.2]	[1.17]	[3.2]	[0.75]
Adj. R^2	0.05		0.08				0.06		

Table 6: Fama-MacBeth estimates for bond factor models. This table reports Fama-MacBeth regression results for three bond factor models. The first model, "CAPMB," is the bond market CAPM where the bond market factor (MKTB) serves as the sole factor. The second model, "Modified BBW," uses the first three factors from Bai, Bali, and Wen (2019)'s four-factor model: the bond market factor, downside risk factor (DRF), and credit risk factor (CRF). The third model, "DEFTERM," is a two-factor model with the default return spread (DEF) and the term return spread (TERM), as proposed by Fama and French (1993). All t-statistics are calculated using Newey-West standard errors with 12 lags and are shown in square brackets beneath the Fama-MacBeth estimates of factor risk premia. Also reported are adjusted cross-sectional R^2 values. The sample period is August 2004 to December 2016 in Panel A and July 1926 to June 2022 in Panel B.

Panel A of Table 6 presents Fama-MacBeth regression results from a relatively short sample period commonly used in the literature. To facilitate comparison, we benchmark our analysis against Dickerson, Mueller, and Robotti (2023) and adopt their sample period, from August 2004 to December 2016.[18] The results from the short sample confirm that none of the examined bond factors produce highly significant risk premia. Both under CAPMB and Modified BBW, the bond market factor shows very weak significance. The insignificant Fama-MacBeth estimates for the downside and credit risk factors imply that these factors are not priced in the cross-section of corporate returns and add little explanatory power beyond the bond market factor. Similarly, the factor risk premia for the default return spread and term return spread under DEFTERM are also insignificant over this sample period.

However, these results change dramatically in Panel B when we extend our sample period. Using a longer sample from July 1926 to June 2022, we find that the bond market factor is now highly significant at the 1% and 5% levels, with t-statistics of 3.84 and 2.53, respectively, under both CAPMB and Modified BBW. Interestingly, the other two factors from Bai, Bali,

[18]Note that Dickerson, Mueller, and Robotti (2023) also test CAPMB and DEFTERM using WRDS data (Internet Appendix Table A6). Although our initial filtering procedure differs slightly from theirs (e.g., we exclude bonds issued by financial firms), the results still closely align with their parameter estimates, t-statistics, and adjusted R^2. We further verify that applying their filtering rules produces even closer results, effectively replicating their values.

and Wen (2019) also show markedly different statistical significance in this extended sample. The downside risk factor becomes statistically significant (t-statistic of 4.98) and so does the credit risk factor (t-statistic of 4.2). In the case of DEFTERM, the default return spread becomes significant, though the term premium remains insignificant. These results suggest that, with the enhanced statistical power provided by our longer sample, multiple bond factors are now clearly priced in the cross-section of corporate bond returns, even after controlling for the bond market factor.

5.3 Traded stock factors

Table 7 turns to three models with stock factors. The first model, "CAPMS," is the classic CAPM with the stock market factor (MKTS). The second model, "FF3," is the Fama-French three-factor model with the stock market factor, size factor (SMB), and value factor (HML). When we run Fama-MacBeth regressions for these models over the short sample period (Panel A), all of these stock factors generate insignificant factor risk premia and thus appear to be unpriced. In the third model, "MKTB vs. MKTS," we examine a two-factor structure where the bond market factor and the stock market factor are juxtaposed. While the risk premium for the bond market factor is weakly significant, the stock market factor is insignificant. These findings resonate with van Binsbergen, Nozawa, and Schwert (2024), who document the failure of stock factor models in explaining corporate bond returns (unless the returns are duration-adjusted), and with Dickerson, Mueller, and Robotti (2023), who demonstrate significant challenges in developing a joint factor structure between the stock and bond markets.

Panel B of Table 7 again shows that these outcomes are driven by the relatively short sample period. Examined over a longer horizon with greater statistical power, not only is the stock market factor significantly priced under all three stock factor models, but also the size and value factors in the Fama-French three-factor model are significantly priced beyond the stock market factor. When both market factors are entered into the same model, neither is driven out by the other, with both remaining highly significant. These findings suggest that

	CAPMS		FF3				MKTB vs. MKTS		
	γ_0	γ_{MKTS}	γ_0	γ_{MKTS}	γ_{SMB}	γ_{HML}	γ_0	γ_{MKTB}	γ_{MKTS}
Panel A: Subsample	07/2002-12/2016		07/2002-12/2016				07/2002-12/2016		
Estimate	0.31	0.93	0.01	1.01	1.15	-2.84	0.00	0.49	0.81
	[3.14]	[1.07]	[0.04]	[1.1]	[0.37]	[-1.67]	[0.01]	[1.41]	[0.76]
Adj. R^2	0.06		0.07				0.09		
Panel B: Full sample	07/1926-06/2022		07/1926-06/2022				07/1926-06/2022		
Estimate	0.21	1.37	-0.05	0.64	2.47	1.42	0.14	0.40	2.71
	[3.18]	[4.49]	[-0.83]	[2.32]	[5.94]	[4.19]	[2.16]	[3.72]	[4.91]
Adj. R^2	0.04		0.06				0.07		

Table 7: Fama-MacBeth estimates for stock factor models. This table reports Fama-MacBeth regression results for three stock factor models. The first model, "CAPMS," is the classic CAPM with the stock market factor (MKTS). The second model, "FF3," is the Fama-French three-factor model with the stock market factor, size factor (SMB), and value factor (HML). The third model, "MKTB vs. MKTS," is a two-factor structure where the bond market factor and the stock market factor are juxtaposed. All t-statistics are calculated using Newey-West standard errors with 12 lags and are shown in square brackets beneath the Fama-MacBeth estimates of factor risk premia. Also reported are adjusted cross-sectional R^2 values. The sample period is August 2004 to December 2016 in Panel A and July 1926 to June 2022 in Panel B.

the previously insignificant risk premia from stock factor models are likely due to insufficient sample lengths. With more statistical power, our corporate bond database offers potential for investigating common factor structures both within corporate bond returns and across stock and bond returns.

5.4 Nontraded factors

Tables 6 and 7 have so far investigated traded factors, which represent portfolio returns reflecting specific sources of systematic risk. Table 8 now considers three prominent nontraded factors that are known to explain expected stock returns: (i) the intermediary factor (CPTL) of He, Kelly, and Manela (2017), which captures shocks to financial intermediaries' capital ratio; (ii) the macro uncertainty factor (UNC) of Jurado, Ludvigson, and Ng (2015), estimated from a wide range of macroeconomic and financial time series; and (iii) the tail risk factor (TAIL) of Kelly and Jiang (2014), extracted from the cross-section of stock returns through a power law estimator. Not surprisingly, these factors are all highly cyclical. We therefore test a two-factor structure where each factor is controlled by the bond market factor to ensure it has incremental explanatory power. The short sample in Panel A spans August 2004 to December 2016. The long sample in Panel B ends in June 2022, with different starting points across models: the intermediary factor begins in January 1970, the macro uncertainty factor in August 1960, and the tail risk factor in August 1926.[19]

	Intermediary			Macro uncertainty			Tail risk		
	γ_0	γ_{MKTB}	γ_{CPTL}	γ_0	γ_{MKTB}	γ_{UNC}	γ_0	γ_{MKTB}	γ_{TAIL}
Panel A: Short sample	07/2002-12/2016			07/2002-12/2016			07/2002-12/2016		
Estimate	0.12	0.34	1.91	0.08	0.35	-0.95	0.18	0.38	0.11
	[0.73]	[1.03]	[1.07]	[0.54]	[1.32]	[-0.83]	[1.16]	[1.13]	[1.66]
Adj. R^2	0.09			0.06			0.06		
Panel B: Long sample	01/1970-06/2022			08/1960-06/2022			08/1926-06/2022		
Estimate	0.15	0.28	3.53	0.08	0.19	-1.96	0.25	0.34	-0.05
	[1.61]	[1.96]	[3.51]	[1.09]	[1.8]	[-4.44]	[3.59]	[3.66]	[-3.58]
Adj. R^2	0.06			0.05			0.06		

Table 8: Fama-MacBeth estimates for nontraded factor models. This table reports Fama-MacBeth regression results for nontraded factor models. We consider three prominent nontraded factors, controlling for the bond market factor: (i) the intermediary factor (CPTL) of He, Kelly, and Manela (2017); (ii) the macro uncertainty factor (UNC) of Jurado, Ludvigson, and Ng (2015); and (iii) the tail risk factor (TAIL) of Kelly and Jiang (2014). All t-statistics are calculated using Newey-West standard errors with 12 lags and are shown in square brackets beneath the Fama-MacBeth estimates of factor risk premia. Also reported are adjusted cross-sectional R^2 values. The short sample in Panel A spans August 2004 to December 2016. The long sample in Panel B ends in June 2022, with different starting points across models: the intermediary factor begins in January 1970, the macro uncertainty factor in August 1960, and the tail risk factor in August 1926.

While Panel A finds that the factor risk premia are statistically insignificant in the short sample period, Panel B demonstrates that the results change drastically when the sample is expanded. In the long sample, the Fama-MacBeth estimates for the three nontraded factors are not only highly significant but also have the "correct" signs, consistent with economic intuition. Like traded factors, the intermediary factor is priced positively in the cross-section, as a positive innovation to financial intermediaries' capital ratio is a "good" shock that lowers investors' marginal utility (or equivalently, the stochastic discount factor). This positive price of risk leads to a positive γ_{CPTL}. In contrast, an increase in the macro uncertainty index or the tail risk index is typically a "bad" shock, coinciding with worse economic conditions and higher investors' marginal utility. Hence, the macro uncertainty and tail risk factors should be priced negatively in the cross-section, resulting in negative values for γ_{UNC} and γ_{TAIL}. This is indeed the case in Panel B where we estimate the models with the longer time series from our corporate bond database.[20]

[19] We download the intermediary factor from Zhiguo He's website, the macro uncertainty factor from Sydney Ludvigson's website, and the tail risk measure from Amit Goyal's website.

[20] In addition to tail risk, we find similar results for the stock volatility factor (SVAR) of Guo (2006) and the

As a word of caution, it is not our goal to identify the best factor model for corporate bonds, as this lies beyond the scope of our paper. We do not aim to determine which factor model explains the data most effectively either. Rather, our exercises in this section demonstrate that the lack of statistical significance in existing studies likely stems from their reliance on short samples.

6 Conclusion

This paper introduces a novel historical database of corporate bonds spanning 128 years, significantly broadening the scope of research on corporate bonds. By integrating hand-collected archival records with contemporary datasets, our database offers a unique opportunity to examine the corporate bond market through multiple economic cycles, including periods of severe financial distress and recovery such as the Great Depression. The sheer breadth and granularity of these data make it possible to investigate asset pricing questions that were previously constrained by data limitations.

As one of our main results, we find that multiple traded bond and stock factors as well as nontraded factors – previously found not to be priced in shorter samples – are actually significantly priced in the cross-section of corporate bonds. These results address critical issues in the asset pricing literature. Using our long sample, the cross-section of corporate bond returns becomes less puzzling: economically plausible bond factors are indeed priced, stock and bond returns appear to share a common factor structure, and key macro variables exhibit explanatory power, consistent with existing theories. These findings suggest that we should be cautious about generalizing conclusions drawn from short samples and highlight the value of our database in deepening our understanding of priced risks in the corporate bond market.

We believe that our corporate bond database can serve as a foundation for future studies, filling many crucial gaps in the literature. For instance, our database, linked with CRSP,

average correlation risk factor (AVGCOR) of Pollet and Wilson (2010). We thank Amit Goyal for providing up-to-date time series of these and many other variables on his website. See Goyal, Welch, and Zafirov (2024) for a comprehensive list of variables he maintains.

facilitates a joint analysis of stocks and corporate bonds issued by the same firms, opening up new opportunities to study the interactions between these two asset classes. This integrated perspective on corporate securities can potentially reveal how market conditions and firm fundamentals affect financing choices across different classes of securities. While this is beyond the scope of the current paper, we plan to examine these issues in future work.

References

Bai, Jennie, Turan G. Bali, and Quan Wen, 2019, RETRACTED: Common risk factors in the cross-section of corporate bond returns, *Journal of Financial Economics* 131, 619–642.

Bali, Turan G, Avanidhar Subrahmanyam, and Quan Wen, 2021, Long-term reversals in the corporate bond market, *Journal of Financial Economics* 139, 656–677.

Chambers, David, Elroy Dimson, Antti Ilmanen, and Paul Rintamäki, 2024, Long-Run Asset Returns, *Annual Review of Financial Economics* 16, 435–458.

Choi, Jaewon, and Yongjun Kim, 2018, Anomalies and market (dis) integration, *Journal of Monetary Economics* 100, 16–34.

Chordia, Tarun, Amit Goyal, Yoshio Nozawa, Avanidhar Subrahmanyam, and Qing Tong, 2017, Are capital market anomalies common to equity and corporate bond markets? An empirical investigation, *Journal of Financial and Quantitative Analysis* 52, 1301–1342.

Chung, Kee H, Junbo Wang, and Chunchi Wu, 2019, Volatility and the cross-section of corporate bond returns, *Journal of Financial Economics* 133, 397–417.

Dick-Nielsen, Jens, Peter Feldhütter, Lasse Heje Pedersen, and Christian Stolborg, 2024, Corporate bond factors: Replication failures and a new framework, *Available at SSRN* 4586652.

Dickerson, Alexander, Philippe Mueller, and Cesare Robotti, 2023, Priced risk in corporate bonds, *Journal of Financial Economics* 150, 103707.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2009, Triumph of the optimists: 101 years of global investment returns, in *Triumph of the Optimists* (Princeton University Press,).

Edwards, Sebastian, 2018, American default: The untold story of FDR, the Supreme Court, and the battle over gold, .

Edwards, Sebastian, Francis A Longstaff, and Alvaro Garcia Marin, 2015, The US debt restructuring of 1933: Consequences and lessons, National Bureau of Economic Research.

Elkamhi, Redouane, Chanik Jo, and Yoshio Nozawa, 2024, A one-factor model of corporate bond premia, *Management Science* 70, 1875–1900.

Fama, Eugene F., and Kenneth R. French, 1992, The cross-section of expected returns, *Journal of Finance* 47, 427–465.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on bonds and stocks, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F, and Kenneth R French, 2004, The capital asset pricing model: Theory and evidence, *Journal of economic perspectives* 18, 25–46.

Fama, Eugene F, and James D MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* pp. 607–636.

Favara, Giovanni, Simon Gilchrist, Kurt Lewis, and Egon Zakrajšek, 2016a, Recession risk and the excess bond premium, *FEDS Notes*, Board of Governors of the Federal Reserve System.

Favara, Giovanni, Simon Gilchrist, Kurt Lewis, and Egon Zakrajšek, 2016b, Updating the recession risk and the excess bond premium, *FEDS Notes*, Board of Governors of the Federal Reserve System.

Gebhardt, William R., Soeren Hvidkjaer, and Bhaskaran Swaminathan, 2005, The cross-section of expected corporate bond returns: Betas or characteristics?, *Journal of Financial Economics* 75, 85–114.

Gilchrist, Simon, and Egon Zakrajšek, 2012, Credit spreads and business cycle fluctuations, *American Economic Review* 102, 1692–1720.

Goldberg, Jonathan, and Yoshio Nozawa, 2021, Liquidity supply in the corporate bond market, *The Journal of Finance* 76, 755–796.

Gomes, Joao F, Mete Kilic, and Sebastien Plante, 2020, Leverage risk and investment: The case of gold clauses in the 1930s, Working paper.

Goyal, Amit, Ivo Welch, and Athanasse Zafirov, 2024, A comprehensive 2022 look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 37, 3490–3557.

Greenwood, Robin, and Samuel G Hanson, 2013, Issuer quality and corporate bond returns, *Review of Financial Studies* 26, 1483–1525.

Guo, Hui, 2006, On the out-of-sample predictability of stock market returns, *The Journal of Business* 79, 645–670.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

Ibbotson, Roger G, and Rex A Sinquefield, 1976, Stocks, bonds, bills, and inflation: year-by-year historical returns (1926-1974), *The Journal of Business* 49, 11–47.

Jensen, Michael C, Fischer Black, and Myron S Scholes, 1972, The capital asset pricing model: Some empirical tests, .

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, 2019, The total risk premium puzzle, Working paper, National Bureau of Economic Research.

Jostova, Gergana, Stanislava Nikolova, Alexander Philipov, and Christof W. Stahel, 2013, Momentum in corporate bond returns, *Review of Financial Studies* 26, 1649–1693.

Jostova, Gergana, Stanislava Stas Nikolova, and Alexander Philipov, 2024, Data Uncertainty in Corporate Bonds, *Alexander, Data Uncertainty in Corporate Bonds (May 24, 2024)*.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.

Kelly, Bryan, and Hao Jiang, 2014, Tail Risk and Asset Prices, *Review of Financial Studies* pp. –.

Kelly, Bryan, Diogo Palhares, and Seth Pruitt, 2023, Modeling corporate bond returns, *The Journal of Finance* 78, 1967–2008.

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

López-Salido, David, Jeremy C Stein, and Egon Zakrajšek, 2017, Credit-market sentiment and the business cycle, *The Quarterly Journal of Economics* 132, 1373–1426.

Pollet, Joshua M, and Mungo Wilson, 2010, Average correlation and stock market returns, *Journal of Financial Economics* 96, 364–380.

Siegel, Jeremy J, 1992, The equity premium: Stock and bond returns since 1802, *Financial Analysts Journal* 48, 28–38.

Siegel, Jeremy J., 1994, *Stocks for the long run: a guide to selecting markets for long-term growth*. (Irwin Burr Ridge, IL).

van Binsbergen, Jules H, Yoshio Nozawa, and Michael Schwert, 2024, Duration-based valuation of corporate bonds, *The Review of Financial Studies* 38, 158–191.

Van Mencxel, Kevin, Jan Annaert, and Marc Deloof, 2024, Bonds for the long run? The rate of return on corporate bonds in Belgium, 1838–1939, *The Economic History Review*.