Common risk factors in the cross-section of corporate bond returns

Jennie Bai, Turan G. Bali, Quan Wen .

Journal of Financial Economics,2019

解读者：屠雪永

2021.04.20

Outline

- Introduction
- Research design
  - Data and variable definitions
- Empirical study
  - Portfolio-level analysis
  - Bond-level Fama-MacBeth regressions
- Conclusion

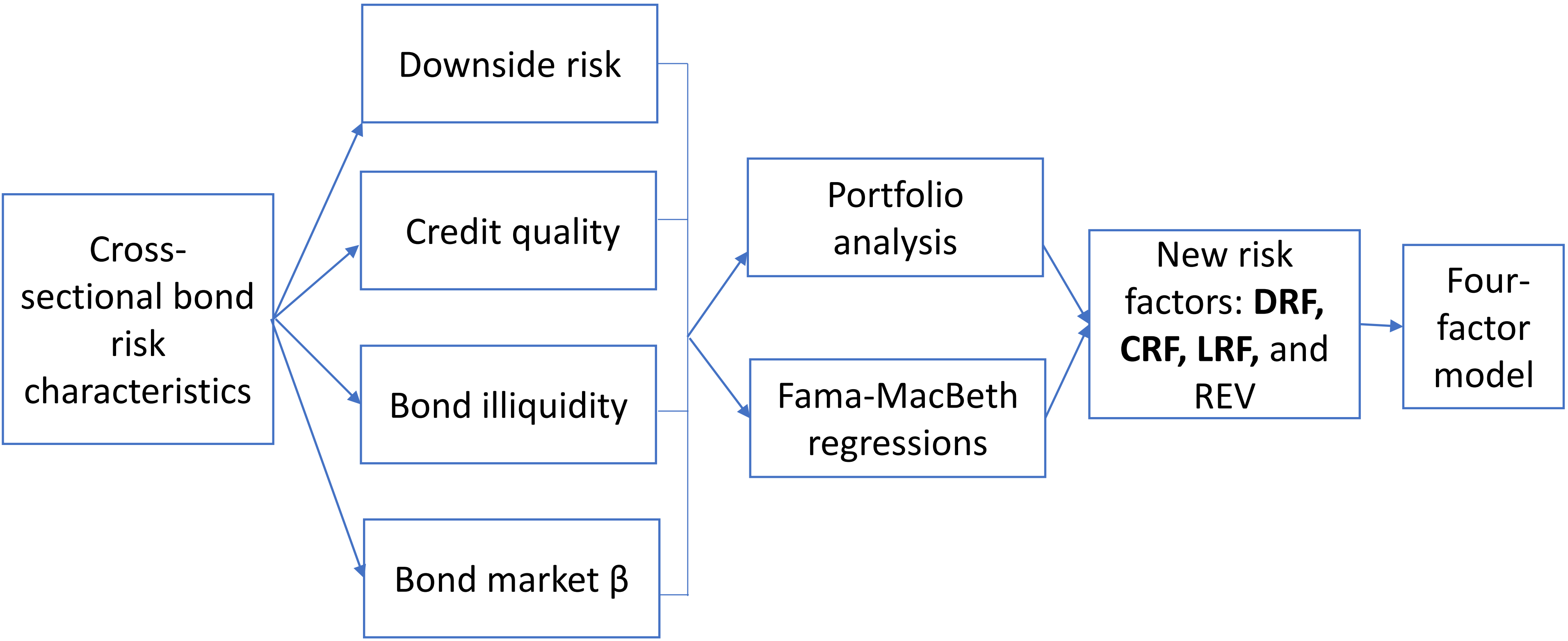
1. Introduction-- Motivation

- An extensive literature examines the cross-sectional determinants of stock returns.
- There is, however, surprisingly little research on the common risk factors that explain the cross-section of corporate bond returns.
- This paper aims to fill this gap by identifying common risk factors that predict the cross-sectional differences in corporate bonds.

1. Introduction-- Motivation

- **Size**
  - total amount   \$19 trillion/\$12 trillion (2016)
  - issuance   \$265 billion / \$1.3 trillion (since 2010)
- **Factors**
  - stock market factors
  - bond market factors
- **Features**
  - downside risk
  - default risk (credit risk)
  - liquidity risk(investors)

1. Introduction-- Framework



1. Introduction-- Contribution

- The foremost contribution is to identify bond-implied **new risk factors** that significantly predict the cross-sectional variation in future bond returns.

- The second contribution of this paper is to demonstrate the empirical performance of **downside risk** in predicting the cross-sectional differences in future returns of corporate bonds.

2. Data-- Corporate bond data

- Database :the enhanced version of the TRACE for July 2002 to December 2016:
  - intraday observations on price, trading volume, and buy and sell indicators
- Merged with the Mergent fixed income securities database(FISD) :
  - offering amount, offering date,maturity date, coupon rate, coupon type, interest payment frequency, bond type, bond rating, bond option features, and issuer information
- Expand the TRACE data by including alternative bond datasets, mainly those containing quoted prices, for a longer sample period starting from January 1977. For this longer sample, we construct downside risk factor and credit risk factor

2. Data-- Corporate bond data

- For TRACE intraday data, we adopt the following filtering criteria:
  - 1. Remove bonds that are not listed or traded in the US public market
  - 2. Remove bonds that are structured notes, mortgage backed or asset backed, agency backed, or equity linked
  - 3. Remove convertible bonds since this option feature distorts the return calculation except callable bonds
  - 4. Remove bonds that trade under $5 or above $1000.
  - 5. Remove bonds that have a floating coupon rate, which means the sample comprises only bonds with a fixed or zero coupon.

2. Data-- Corporate bond data

- 6. Remove bonds that have less than one year to maturity.
- 7. For intraday data, we also eliminate bond transactions that are labeled as when-issued, locked-in, or have special sales conditions, and that have more than a two day settlement.
- 8. Remove transaction records that are canceled and adjust records that are subsequently corrected or reversed.
- 9. Remove transaction records that have trading volume less than $10,000

2. Data-- Corporate bond return

- The monthly corporate bond return at time t is computed as

$$r_{i,t} = \frac{P_{i,t} + AI_{i,t} + C_{i,t}}{P_{i,t-1} + AI_{i,t-1}} - 1, \qquad (1)$$

- Bond i’s excess return

$$R_{i,t} = r_{i,t} - r_{f,t}$$

- Where $r_{f,t}$ is the risk-free rate proxied by the one-month Treasury bill rate.

2. Data-- Corporate bond data

- We first calculate the **daily clean price** as the trading volume-weighted average of intraday prices
- convert the bond prices from daily to monthly frequency:
    - **(i) from the end of month t − 1 to the end of month t**
    - (ii) from the beginning of month t to the end of month t.
    - the end (beginning) of month refers to the last (first) five trading days within each month

- **38,957 bonds issued by 4079 unique firms,** approximately 7147 bonds per month over the whole sample

2. Data-- Cross-sectional bond risk characteristics

- The two markets are integrated:
  - First, both bonds and stocks are contingent claims on the value of the same underlying assets.
  - Second, the expected default loss of corporate bonds changes with equity price.

2. Data-- Cross-sectional bond risk characteristics

The corporate bond market has its own unique features:

- First, **credit risk** is particularly important in determining corporate bond returns.
- Second, bondholders are more sensitive to **downside risk** than stockholders.
- Third, the corporate bond market is less **liquid** than the equity market.
- Fourth, corporate bond market participants have been dominated by institutional investors**.**
- Finally, there is some evidence that shows the discrepancy in return premiums between equity and corporate bond markets.

2. Data-- Downside risk

- VaR, which determines how much **the value of an asset** could decline over a **given period of time** with **a given probability** as a result of changes in market rates or prices.
- Our proxy for downside risk, 5% VaR, is the second lowest monthly return observation over the past 36 months.

- Expected shortfall (ES), defined as the conditional expectation of loss given that the loss is beyond the VaR level.
- 10% VaR and 10% ES for rubustness check

2. Data-- Credit quality

- We collect bond-level rating information from **Mergent Fixed Income Securities Database (FISD)** historical ratings.

- All ratings are assigned a number to facilitate the analysis;
  Investment-grade bonds have ratings from 1 (AAA) to 10 (BBB-).
  Noninvestment-grade bonds have ratings above 10.

- We determine a bond’s rating as the average of ratings provided by Standard & Poor (S&P) and Moody’s when both are available or one if only one is available.

2. Data-- Bond illiquidity

- We follow Bao, Pan, and Wang (2011) to construct bond level illiquidity measure, ILLIQ, which aims to extract the transitory component from bond price. Specifically, the log price change for bond i on day d of month t:

$$\Delta p_{itd} = p_{itd} - p_{itd-1}$$

- Then, ILLIQ is defined as

$$ILLIQ = -Cov_t(\Delta p_{itd}, \Delta p_{itd+1}). \qquad (2)$$

2. Data-- Bond market β

We estimate the bond market beta, $\boldsymbol{\beta^{Bond}}$**,** for each bond from the time-series regressions of individual bond excess returns on the bond market excess returns using a **36-month rolling** window.

•

2. Data-- Summary statistics

Panel A: Cross-sectional statistics over the sample period of July 2002–December 2016

	N	Mean	Median	SD	Percentiles 1st	5th	25th	75th	95th	99th
Bond return (%)	1,243,543	0.68	0.50	3.13	−7.46	−3.66	−0.68	1.86	5.59	10.33
Rating	1,243,543	8.32	7.65	4.05	1.56	2.25	5.52	10.35	16.30	19.09
Time to maturity (maturity, year)	1,243,543	9.49	6.60	8.26	1.11	1.51	3.59	12.81	26.69	31.63
Amount out (size, $million)	1,243,543	393.73	269.59	478.63	1.60	5.17	76.99	504.15	1353.24	2480.32
Downside risk (5% VaR)	579,333	5.84	4.08	5.78	0.70	1.17	2.46	6.96	16.75	29.42
Illiquidity (ILLIQ)	977,011	2.14	0.46	5.17	−1.17	−0.23	0.07	1.99	10.16	24.13
Bond market beta (β^{bond})	584,223	1.12	1.01	1.15	−0.24	0.15	0.58	1.67	3.72	5.38

Panel B: Average cross-sectional correlations

	Rating	Maturity	Size	VaR	ILLIQ	β^{Bond}
Rating	1	−0.138	−0.021	0.383	0.117	0.089
Maturity		1	−0.042	0.171	0.106	0.356
Size			1	−0.108	−0.160	0.076
VaR				1	0.323	0.195
ILLIQ					1	0.098
β^{Bond}						1

3.Downside risk and expected corporate bond returns

- **Normality test for corporate bond returns**

Panel A: Time-series distribution of all corporate bond returns

	Volatility	Skewness Positive	Skewness Negative	Kurtosis Positive	Kurtosis Negative	Normality JB-stat
Total # of bonds	38,957	19,548	19,409	26,493	12,464	38,957
% of bonds significant	84.57%	48.02%	99.45%	67.65%	52.60%	79.91%
Median p-value	0.00	0.00	0.00	0.00	0.00	0.00
% of bonds insignificant	15.43%	51.98%	0.55%	32.35%	47.40%	20.09%

Panel B: Cross-sectional distribution of monthly corporate bond returns

	Volatility	Skewness Positive	Skewness Negative	Kurtosis Positive	Kurtosis Negative	Normality JB-stat
Total # of months	150	118	32	150	0	150
# of months significant	150	118	32	150	0	150
Median p-value	0.00	0.00	0.00	0.00	0.00	0.00
# of months insignificant	0	0	0	0	0	0

- Bond returns is skewed, peaked around the mode, and has fat tails, downside risk—defined as a nonlinear function of volatility, skewness, and kurtosis

3. **Univariate portfolios** of corporate bonds sorted by **downside risk**

Quintiles	Average VaR	Average return	Five-factor stock alpha	Five-factor bond alpha	Ten-factor alpha	Average portfolio characteristics				
						β^{Bond}	ILLIQ	Rating	Maturity	Size
Low VaR	1.59	0.21	0.19	0.03	0.03	0.55	0.57	7.02	4.43	0.56
		(1.10)	(1.28)	(1.05)	(1.09)					
2	2.95	0.34	0.30	0.04	0.05	0.82	1.15	7.69	7.07	0.46
		(2.99)	(2.78)	(1.17)	(1.35)					
3	4.38	0.44	0.37	0.04	0.04	1.06	1.82	7.91	10.39	0.43
		(2.77)	(2.58)	(0.65)	(0.79)					
4	6.71	0.62	0.54	0.17	0.19	1.44	2.72	8.64	13.26	0.41
		(3.02)	(3.32)	(1.82)	(1.98)					
High VaR	15.72	1.20	0.99	0.81	0.75	2.52	5.20	12.16	12.15	0.34
		(4.18)	(4.41)	(4.32)	(3.16)					
High - Low	14.13***	0.99***	0.79***	0.78***	0.72***					
Return/Alpha diff.	(9.94)	(3.95)	(3.82)	(3.90)	(2.82)					

- The positive relation between VaR and the cross-section of bond returns
- Bonds with high downside risk have a higher market beta, lower liquidity, higher credit risk, longer time to maturity, and smaller size.

3. **Bivariate portfolios** of corporate bonds sorted by **downside risk controlling for bond characteristics.**

	Panel A: Controlling for credit rating			Panel B: Controlling for maturity			
	All bonds	Investment grade	Noninvestment grade	All bonds	Short maturity	Medium maturity	Long maturity
VaR,1	0.04	0.02	0.32	−0.05	−0.08	−0.06	0.01
	(1.07)	(0.96)	(3.01)	(−1.19)	(−1.22)	(−1.34)	(0.44)
VaR,2	0.11	0.05	0.31	−0.02	−0.14	−0.01	0.07
	(3.75)	(2.25)	(3.15)	(−0.57)	(−2.47)	(−0.35)	(3.06)
VaR,3	0.12	0.03	0.43	0.12	0.08	0.11	0.18
	(3.11)	(1.23)	(3.68)	(3.23)	(1.37)	(2.03)	(4.73)
VaR,4	0.18	−0.03	0.42	0.20	0.16	0.23	0.27
	(3.02)	(−1.16)	(2.61)	(2.67)	(1.42)	(1.71)	(3.77)
VaR,5	0.51	0.40	1.10	0.86	0.56	0.74	1.00
	(3.41)	(1.39)	(4.46)	(4.36)	(2.78)	(2.91)	(4.83)
VaR,5 - VaR,1 Return/Alpha diff.	0.46**	0.38**	0.78***	0.91***	0.64***	0.81***	0.99***
	(2.60)	(2.46)	(3.38)	(4.10)	(2.66)	(2.88)	(4.59)

	Panel C: Controlling for size			Panel D: Controlling for illiquidity		
	All bonds	Small bonds	Large bonds	All bonds	Investment grade	Noninvestment grade
VaR,1	0.06	0.04	0.05	−0.01	−0.01	0.13
	(0.96)	(0.55)	(0.96)	(−0.20)	(−0.28)	(1.15)
VaR,2	0.08	0.18	0.06	0.07	0.02	0.30
	(1.20)	(1.94)	(1.57)	(2.69)	(0.89)	(2.54)
VaR,3	0.19	0.38	0.08	0.09	0.03	0.39
	(4.16)	(4.09)	(1.63)	(1.98)	(1.41)	(3.03)
VaR,4	0.26	0.49	0.10	0.19	0.01	0.65
	(3.10)	(3.05)	(1.56)	(2.50)	(0.45)	(3.43)
VaR,5	0.71	0.83	0.65	0.72	0.37	1.29
	(4.17)	(2.77)	(3.50)	(4.26)	(2.03)	(4.46)
VaR,5 - VaR,1 Return/Alpha diff.	0.65***	0.79**	0.60***	0.72***	0.38**	1.16***
	(3.48)	(2.37)	(3.08)	(3.90)	(2.48)	(3.55)

3. Bond-level Fama-MacBeth regressions.

	Intercept	5% VaR	Rating	ILLIQ	β^{Bond}	β^{DEF}	β^{TERM}	Maturity	Size	REV	Adj. R^2
(1)	−0.011	**0.064**									0.086
	(−0.10)	(4.88)									
(2)	0.127	**0.052**				−0.006	0.002	−0.002	−0.012	**−0.122**	0.173
	(0.96)	(4.36)				(−1.03)	(0.23)	(−0.46)	(−0.84)	(−9.24)	
(3)	−0.182		**0.068**								0.054
	(−1.32)		(3.84)								
(4)	−0.130		**0.064**			−0.008	0.018	**0.015**	−0.001	**−0.119**	0.155
	(−1.23)		(2.84)			(−1.46)	(1.28)	(2.12)	(−1.00)	(−9.36)	
(5)	0.463			**0.081**							0.028
	(3.41)			(6.45)							
(6)	0.304			**0.066**		−0.007	0.041	0.007	0.030	**−0.079**	0.152
	(2.68)			(6.32)		(−0.90)	(1.37)	(1.17)	(0.99)	(−5.29)	
(7)	0.209				**0.486**						0.055
	(1.72)				(3.15)						
(8)	0.224				**0.318**	**−0.023**	0.026	0.004	−0.053	**−0.069**	0.156
	(2.50)				(2.14)	(−2.76)	(1.63)	(0.72)	(−1.13)	(−3.27)	
(9)	−0.195	**0.111**	0.031	**0.047**	−0.097						0.144
	(−1.37)	(5.29)	(1.50)	(6.22)	(−0.94)						
(10)	−0.178	**0.106**	0.030	**0.041**	−0.097	−0.002	0.003	0.002	0.000	**−0.132**	0.217
	(−1.55)	(4.72)	(1.48)	(5.25)	(−0.95)	(−0.31)	(0.31)	(0.30)	(3.22)	(−8.51)	

$$R_{i,t+1} = \lambda_{0,t} + \lambda_{1,t} VaR_{i,t} + \lambda_{2,t} Rating_{i,t} + \lambda_{3,t} ILLIQ_{i,t} + \lambda_{4,t} \beta_{i,t}^{Bond} + \sum_{k=1}^{K} \lambda_{i,k,t} Control_{i,k,t} + \epsilon_{i,t+1}, \qquad (3)$$

3. The source of downside risk premium

Panel A: Trivariate dependent-sort portfolios by VOL, SKEW, and KURT

	Average return	Ten-factor alpha		Average return	Ten-factor alpha		Average return	Ten-factor alpha
Low VOL	0.27	0.03	Low SKEW	0.78	0.45	Low KURT	0.49	0.01
	(1.29)	(0.91)		(4.56)	(3.65)		(3.03)	(0.10)
2	0.47	0.05	2	0.65	0.26	2	0.53	0.10
	(2.05)	(0.92)		(3.11)	(2.29)		(3.35)	(1.73)
High VOL	0.91	0.37	High SKEW	0.53	0.23	High KURT	0.55	0.21
	(3.66)	(2.56)		(2.79)	(2.39)		(3.89)	(4.23)
High - Low	0.64***	0.34**	High - Low	−0.25**	−0.22**	High - Low	0.06	0.20**
t-stat	(3.32)	(2.48)	*t*-stat	(−2.47)	(−2.35)	*t*-stat	(0.81)	(2.34)

Panel B: Cross-sectional regressions with VOL, SKEW, and KURT

	Intercept	VOL	SKEW	KURT	Rating	Maturity	Size	β^{DEF}	β^{TERM}	REV	Adj. R^2
(1)	0.231	**0.011**									0.076
	(2.03)	(2.92)									
(2)	0.128	**0.012**			0.025	0.004	−0.011	−0.001	−0.010	**−0.126**	0.180
	(1.15)	(2.66)			(1.22)	(0.84)	(−0.55)	(−0.10)	(−1.06)	(−9.28)	
(3)	0.206	**0.011**	**−0.128**								0.083
	(1.80)	(2.96)	(−2.81)								
(4)	0.127	**0.012**	**−0.099**		0.025	0.005	−0.014	−0.000	−0.010	**−0.125**	0.184
	(1.16)	(2.74)	(−2.88)		(1.21)	(0.85)	(−0.73)	(−0.09)	(−1.01)	(−9.24)	
(5)	0.375			**0.047**							0.012
	(2.26)			(2.73)							
(6)	0.071			**0.016**	**0.052**	**0.012**	−0.034	−0.005	0.013	**−0.108**	0.159
	(0.54)			(2.34)	(1.97)	(2.02)	(−1.13)	(−0.73)	(1.07)	(−7.95)	
(7)	0.214	**0.011**	**−0.102**	−0.001							0.087
	(1.81)	(2.89)	(−2.43)	(−0.05)							
(8)	0.129	**0.012**	**−0.081**	0.002	0.025	0.005	−0.016	−0.000	−0.010	**−0.125**	0.186
	(1.16)	(2.57)	(−2.53)	(0.26)	(1.27)	(0.94)	(−0.86)	(−0.07)	(−1.05)	(−9.36)	

4. **Common risk factors in the corporate bond market**

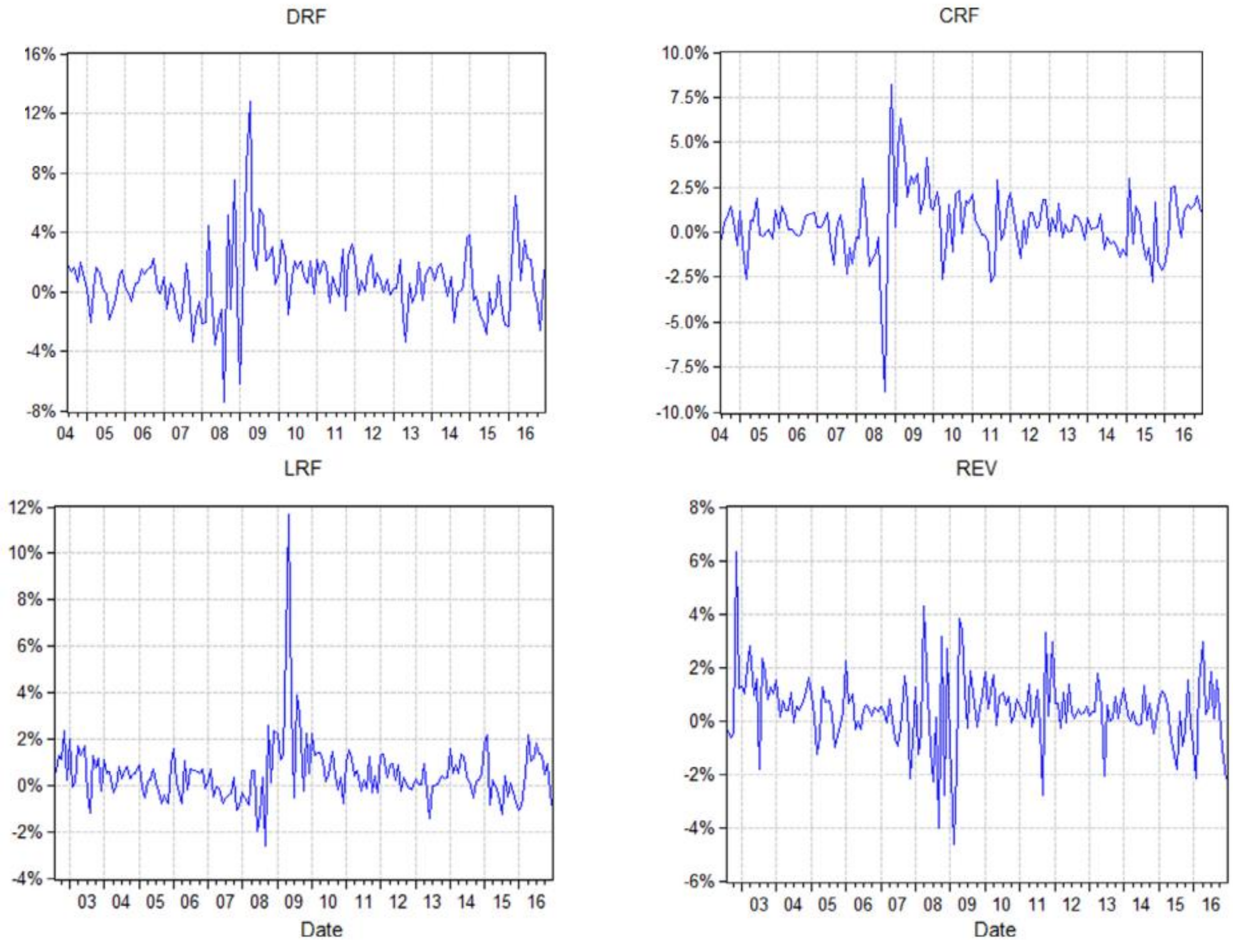
- MKT^{Bond} is the corporate bond **market excess return** constructed using the value-weighted average return of all corporate bonds in the sample (in excess of onemonth T-bill rate).

- Downside risk factor (*DRF*) is constructed by **independently** sorting corporate bonds into **5 × 5** quintiles based on the 5% VaR and credit rating.
  - *DRF* is the value-weighted average return difference between the **highest-VaR portfolio minus the lowest-VaR portfolio** within each rating portfolio. (also obtain CRF_{VaR})

- Liquidity risk factor (*LRF*)
- Return reversal factor (*REV*)

- Credit risk factor (CRF) is the **average of the CRF** obtained from forming the DRF, LRF, and REV, and CRF = 1/3(CRF_{VaR} + CRF_{ILLIQ}+ CRF_{REV})

4. **Common risk factors in the corporate bond market**



The newly proposed DRF, CRF, and LRF risk factors generate economically large risk premiums during economic downturns

4. **Common risk factors in the corporate bond market**

Panel A: Summary statistics on the value-weighted bond factors		
	Mean	*t*-stat
MKT^{Bond}	0.39	3.58
Downside risk factor (DRF)	$\boxed{0.70}$	3.60
Credit risk factor (CRF)	0.43	2.78
Liquidity risk factor (LRF)	0.52	5.02
Return reversal factor (REV)	0.41	4.05

Panel B: Factor alpha from the ten-factor model			
	Model 1	Model 2	Model 3
DRF alpha	0.83	0.79	0.80
t-stat	(2.90)	(3.19)	(2.76)
CRF alpha	0.44	0.34	0.35
t-stat	(2.92)	(2.01)	(1.89)
LRF alpha	0.37	0.32	0.32
t-stat	(3.15)	(2.79)	(2.45)
REV alpha	0.48	0.49	0.46
t-stat	(4.10)	(4.46)	(4.74)

- These novel factors capture an important source of common return variation in corporate bonds missing from long-established stock and bond market factors

4. **Are exposures to bond risk factors priced?**

	Intercept	β^{Bond}	β^{DRF}	β^{CRF}	β^{LRF}	β^{REV}	VaR	Rating	ILLIQ	REV	Maturity	Size	Adj. R^2
(1)	0.513	**0.301**	**0.399**	**0.816**	**0.267**	−0.174							0.103
	(3.32)	(2.85)	(2.99)	(4.89)	(2.39)	(−1.27)							
(2)	−0.088	**0.202**	**0.275**	**0.118**	**0.272**	−0.049	**0.158**						0.152
	(−0.97)	(2.34)	(2.78)	(2.23)	(2.31)	(−0.54)	(4.61)						
(3)	−0.237	**0.269**	**0.403**	**0.430**	**0.321**	−0.095		**0.095**					0.128
	(−1.74)	(2.63)	(2.37)	(4.36)	(2.74)	(−0.71)		(3.27)					
(4)	0.315	**0.282**	**0.443**	**0.545**	**0.331**	−0.076			**0.095**				0.136
	(2.94)	(2.66)	(2.82)	(3.86)	(2.33)	(−0.54)			(5.59)				
(5)	0.466	**0.261**	**0.321**	**0.627**	**0.162**	−0.234				**−0.035**			0.125
	(3.90)	(2.74)	(2.69)	(4.01)	(2.22)	(−1.91)				(−3.69)			
(6)	0.411	**0.296**	**0.357**	**0.799**	**0.255**	−0.165					0.009		0.129
	(3.12)	(2.71)	(2.75)	(4.83)	(2.28)	(−1.16)					(1.44)		
(7)	0.975	0.318	**0.434**	**0.737**	**0.305**	−0.132						−0.090	0.112
	(2.31)	(1.28)	(2.92)	(4.73)	(2.36)	(−0.93)						(−1.68)	
(8)	−0.346	0.172	**0.351**	**0.493**	**0.289**	−0.102	**0.126**	0.044	**0.055**	**−0.118**	−0.005	0.007	0.234
	(−2.90)	(1.05)	(2.51)	(2.60)	(2.85)	(−1.17)	(5.60)	(1.11)	(6.81)	(−7.40)	(−0.88)	(0.36)	

$$R_{i,t} = \alpha_{i,t} + \beta_{i,t}^{MKT} \cdot MKT_t^{Bond} + \beta_{i,t}^{Factor} \cdot Factor_t + \epsilon_{i,t}, \qquad (4)$$

- REV is a nonrisk bond characteristic instead of a common risk factor in the bond market.
- The DRF, CRF, and LRF remain significant risk factors along with downside risk, illiquidity, and one month lagged return as significant characteristics

4. Alternative test portfolios

- **25-size/maturity-sorted bond portfolios**

Panel A: Model 1																	
	Alpha (α)						t-statistics						Adj. R^2				
	Short	2	3	4	Long		Short	2	3	4	Long		Short	2	3	4	Long
Small	0.38	0.53	0.58	0.42	0.55	Small	3.44	3.74	3.49	2.63	3.26	Small	0.10	0.09	0.07	0.07	0.09
2	0.31	0.47	0.51	0.52	0.51	2	4.06	4.06	3.60	2.60	3.13	2	0.10	0.10	0.05	0.05	0.07
3	0.24	0.38	0.41	0.43	0.53	3	4.18	4.16	3.25	3.18	3.04	3	0.17	0.13	0.10	0.05	0.01
4	0.23	0.31	0.41	0.37	0.50	4	3.71	3.50	3.15	2.57	2.50	4	0.13	0.08	0.07	0.04	0.02
Big	0.14	0.31	0.40	0.41	0.52	Big	2.30	3.14	2.91	2.70	2.35	Big	0.05	0.04	0.05	0.03	0.02
Average $\lvert\alpha\rvert$	**0.42**											Average R^2	**0.07**				
p-GRS	<0.01																
Panel B: Model 2																	
Average $\lvert\alpha\rvert$	**0.33**											Average R^2	**0.18**				
p-GRS	<0.01																
Panel C: Model 3																	
Average $\lvert\alpha\rvert$	**0.14**											Average R^2	**0.27**				
p-GRS	0.03																
Panel D: Model 4																	
Average $\lvert\alpha\rvert$	**0.04**											Average R^2	**0.56**				
p-GRS	0.06																
Panel E: Model 5																	
Average $\lvert\alpha\rvert$	**0.03**											Average R^2	**0.57**				
p-GRS	0.06																

Model1: MKT^{Stock}, SMB,HML, MOM^{Stock} and LIQ^{Stock}.

Model2: MKT^{Bond} ,the default factor (DEF), the term factor (TERM), MOM^{Bond} and LIQ^{Bond}.

Model3: MKT^{Bond} ,the credit risk factor (CRF), and the bond liquidity risk factor (LRF)

Model4: MKT^{Bond} , the downside risk factor (**DRF**), the credit risk factor (CRF), and the bond liquidity risk factor (LRF)

Model5: MKT^{Bond} , the downside risk factor (DRF), the credit risk factor (CRF),bond liquidity risk factor (LRF) and the return reversal factor (**REV**)

4. **Alternative test portfolios**

- **30 industry-sorted bond portfolios**

Industry #	Industry description	Model 1			Model 2			Model 3			Model 4			Model 5		
		Alpha (α)	$t(\alpha)$	R^2	Alpha (α)	$t(\alpha)$	R^2	Alpha (α)	$t(\alpha)$	R^2	(α)	$t(\alpha)$	R^2	(α)	$t(\alpha)$	R^2
1	Food	0.37	3.33	0.10	0.25	2.42	0.23	0.10	0.94	0.25	0.08	0.79	0.27	0.16	1.51	0.30
2	Beer	0.28	3.19	0.05	0.22	2.59	0.14	0.14	1.69	0.18	0.11	1.43	0.30	0.14	1.66	0.30
3	Tobacco	0.43	2.40	0.08	0.39	2.11	0.03	0.21	1.12	0.10	0.14	0.83	0.21	0.24	1.33	0.23
4	Games	0.79	2.24	0.13	0.69	1.92	0.13	0.11	0.33	0.30	0.02	0.05	0.36	0.05	0.15	0.36
5	Books	0.55	1.86	0.38	0.44	1.40	0.32	−0.22	−0.77	0.46	−0.31	−1.17	0.52	−0.22	−0.79	0.53
6	Household	0.45	2.17	0.10	0.39	1.85	0.09	0.22	1.02	0.14	0.14	0.71	0.26	0.18	0.87	0.26
7	Clothes	0.68	2.24	0.26	0.40	1.43	0.37	−0.18	−0.60	0.36	−0.22	−0.78	0.38	0.00	0.01	0.41
8	Health	0.48	2.69	0.02	0.41	2.32	0.09	0.18	1.04	0.16	0.16	0.90	0.17	0.22	1.23	0.18
9	Chemicals	0.52	2.42	0.35	0.40	1.79	0.29	−0.08	−0.37	0.42	−0.17	−0.88	0.54	−0.05	−0.24	0.55
10	Textiles	0.66	1.59	0.03	0.48	1.13	0.04	0.19	0.44	0.05	0.13	0.31	0.07	0.27	0.61	0.07
24	Paper	0.50	2.08	0.24	0.37	1.52	0.24	−0.12	−0.60	0.51	−0.19	−1.01	0.57	−0.06	−0.29	0.59
25	Transportation	0.54	3.40	0.13	0.44	3.10	0.34	0.14	1.11	0.47	0.10	0.80	0.55	0.17	1.35	0.56
26	Wholesale	0.46	2.59	0.12	0.30	1.86	0.28	0.07	0.47	0.43	0.03	0.20	0.48	0.11	0.72	0.49
27	Retail	0.54	2.36	0.12	0.37	1.64	0.17	0.02	0.09	0.34	−0.04	−0.19	0.39	0.02	0.12	0.39
28	Restaurant	0.40	1.55	0.14	0.28	1.09	0.16	−0.25	−1.14	0.43	−0.32	−1.52	0.49	−0.20	−0.92	0.50
29	Finance	0.43	3.30	0.08	0.36	2.85	0.15	0.03	0.30	0.53	−0.01	−0.08	0.61	−0.01	−0.06	0.60
30	Other	0.73	3.27	0.15	0.55	2.55	0.25	−0.05	−0.30	0.53	−0.07	−0.43	0.54	0.03	0.15	0.55
Average $\|\alpha\|$		**0.55**		**0.13**	**0.41**		**0.18**	**0.15**		**0.31**	**0.14**		**0.37**	**0.12**		**0.38**
p-GRS		< 0.01			< 0.01			0.03				0.03		0.03		

- These results provide supporting evidence for the remarkable performance of the newly proposed factors in predicting the cross-sectional variation in the returns of the 30-industry portfolios of corporate bonds

V. Conclusion

- **Downside risk, credit risk, and liquidity risk** positively predict the cross-sectional variation in future bond returns. And we then introduce novel risk factors based on these prevalent bond risk characteristics.

- We also find a **strong short-term reversal effect** in the cross-section of corporate bond returns but one-month lagged return is a strong nonrisk bond characteristic instead of a common risk factor in the bond market.

Corporate Bond Characteristics

43 corporate bond characteristics, 4 categories

- I: Bond characteristics related to **interest risk or maturity**:
  - Rating, MAT , Size, Age,DUR.
- II: Bond characteristics related to **risk measures** such as downside risk or systematic risk:
  - VaR5, VaR10, ES5, ES10, βBond, βDEF , βTERM, βUNC, βVIX, βLWW , βDRF , βCRF , βLRF , COSKEW **.**
- III: Bond-level **illiquidity and illiquidity risk**:
  - ILLIQ, Roll, TCRoll, P HighLow,P Zeros, P FHT, Amihud, PI Roll, PI FHT, PI HighLow, StdAmihud, PI Lambda, AvgBidAsk, TC IQR, Roundtrip, γPS.
- IV: **Past return characteristics**:
  - REV , MOM6, MOM12, LTR, VOL, SKEW ,KURT , ISKEW .

Bali, Turan G, Amit Goyal, Dashan Huang, Fuwei Jiang, and Quan Wen. “Different Strokes: Return Predictability Across Stocks and Bonds with Machine Learning and Big Data,” n.d., 81.

1. **Credit rating (*Rating*).** We collect bond-level rating information from Mergent FISD historical ratings. All ratings are assigned a number to facilitate the analysis, for example, 1 refers to a AAA rating, 2 refers to AA+, ..., and 21 refers to CCC. Investment-grade bonds have ratings from 1 (AAA) to 10 (BBB−). Non-investment-grade bonds have ratings above 10. A larger number indicates higher credit risk, or lower credit quality. We determine a bond's rating as the average of ratings provided by S&P and Moody's when both are available, or as the rating provided by one of the two rating agencies when only one rating is available.

2. **Time-to-maturity (*MAT*).** The number of years to maturity.

3. **Issuance size (*Size*).** The natural logarithm of bond amount outstanding.

4. **Age (*Age*).** Bond age since the first issuance, in the number of years.

5. **Duration (*DUR*).** A bond's price sensitivity to interest rate changes, measured in years.

6. **Downside risk proxied by the 5% VaR (*VaR5*).** Following Bai, Bali, and Wen (2019), we measure downside risk of corporate bonds using VaR, which determines how much the value of an asset could decline over a given period of time with a given probability as a result of changes in market rates or prices. Our proxy for downside risk, 5% Value-at-Risk (VaR), is based on the lower tail of the empirical return distribution, that is, the second lowest monthly return observation over the past 36 months. We then multiply the original measure by −1 for convenience of interpretation.[16]

7. **Downside risk proxied by the 10% VaR (*VaR10*).** This measure is defined as the fourth lowest monthly return observation over the past 36 months. We then multiply the original measure by −1 for convenience of interpretation.

8. **Downside risk proxied by the 5% Expected Shortfall (*ES5*).** An alternative measure of downside risk, "expected shortfall," is defined as the conditional expectation of loss given that the loss is beyond the VaR level. In our empirical analyses, we use the 5% expected shortfall (ES5) defined as the average of the two lowest monthly return observations over the past 36 months (beyond the 5% VaR threshold).

9. **Downside risk proxied by the 10% Expected Shortfall (*ES10*).** An alternative measure of downside risk, "expected shortfall," is defined as the conditional expectation of loss given that the loss is beyond the VaR level. In our empirical analyses, we use the 10% expected shortfall (ES10) defined as the average of the four lowest monthly return observations over the past 36 months (beyond the 10% VaR threshold).

10. **Illiquidity (*ILLIQ*).** A bond-level illiquidity measure. We follow Bao, Pan, and Wang (2011) to construct the measure, which aims to extract the transitory component from bond price. Specifically, let $\Delta p_{itd} = p_{itd} - p_{itd-1}$ be the log price change for bond i on day d of month t. Then, $ILLIQ$ is defined as

$$ILLIQ = -Cov_t(\Delta p_{itd}, \Delta p_{itd+1}).$$

11. **Roll's daily measure of illiquidity (Roll).** As an alternative measure of bond-level illiquidity using daily bond returns, the Roll (1984) measure is defined as,

$$\text{Roll} = \begin{cases} 2\sqrt{-\text{cov}(r_d, r_{d-1})} & \text{if } \text{cov}(r_d, r_{d-1}) < 0, \\ 0 & \text{otherwise,} \end{cases}$$

where r_d is the corporate bond return on day d.

12. **Roll's intraday measure of illiquidity (TC_Roll).** Following Dick-Nielsen, Feldhütter, and Lando (2012), we employ an intraday version of the Roll (1984) estimator for effective spreads,

$$\text{TC\_Roll} = \begin{cases} 2\sqrt{-\text{cov}(r_i, r_{i-1})} & \text{if } \text{cov}(r_i, r_{i-1}) < 0, \\ 0 & \text{otherwise,} \end{cases}$$

where $r_i = \frac{P_i - P_{i-1}}{P_{i-1}}$ is the return of the ith trade.

13. **High-low spread estimator(P_HighLow).** Following Corwin and Schultz (2012), we use the ratio between the daily high and low prices on consecutive days to approximate bid-ask spreads. With such motivation, their effective spread proxy is defined as

$$\begin{aligned} \text{P\_HighLow} &= \frac{2(e^{\alpha} - 1)}{1 + e^{\alpha}}, \\ \alpha &= \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \\ \beta &= \sum_{j=0}^{1} \left(\ln \left(\frac{H_{t+j}}{L_{t+j}} \right) \right)^2, \\ \gamma &= \left(\ln \left(\frac{H_{t,t+1}}{L_{t,t+1}} \right) \right)^2. \end{aligned}$$

$H_t(L_t)$ is the highest (lowest) transaction price at day t, and $H_{t,t+1}(L_{t,t+1})$ is the highest (lowest) price on two consecutive days t and $t+1$. Again, we take the mean of the daily values in a month to get a monthly spread proxy for each bond.

14. **Illiquidity measure based on zero returns (P_Zeros).** Following Lesmond, Ogden, and Trzcinka (1999), we use the proportion of zero return days as a measure of liquidity. Lesmond, Ogden, and Trzcinka (1999) argue that zero volume days (hence zero return days) are more likely to reflect lower liquidity. We compute their measure on a monthly basis with T as the number of trading days in a month,

$$\text{P\_Zeros} = \frac{\#\text{ of zero return days}}{T},$$

The number of zero return days comprises two parts, the sequential days with no price change hence zero returns, and the days with zero trading volume.

15. **Modified illiquidity measure based on zero returns (P_FHT).** Fong, Holden, and Trzcinka (2017) propose a new bid-ask spread proxy based on the zeros measure in Lesmond, Ogden, and Trzcinka (1999). In their framework, symmetric transaction costs of $S/2$ leads to observed returns of

$$R = \begin{cases} R^* + \frac{S}{2} & \text{if } R^* < -\frac{S}{2}, \\ 0 & \text{if } -\frac{S}{2} < R^* < \frac{S}{2}, \\ R^* - \frac{S}{2} & \text{if } \frac{S}{2} < R^*, \end{cases}$$

where R^* is the unobserved true value return, which they assume to be normally distributed with mean zero and variance σ^2. Hence, they equate the theoretical probability of a zero return with its empirical frequency, measured via P_Zeros. Solving for the spread S, they get

$$\text{P\_FHT} = S = 2 \cdot \sigma \cdot \Phi^{-1}\left(\frac{1 + \text{P\_Zeros}}{2}\right)$$

where Φ^{-1} is the inverse of the cumulative standard normal distribution. We compute a bond's σ for each month and then calculate P_FHT.

$P\_zeros = P(R=0)$ —— 可观测

由于假设 $R^* \sim N(0, \sigma^2)$

则 $P(-\frac{s}{2} < R^* < \frac{s}{2}) = \Phi\left(\frac{\frac{s}{2}-0}{\sigma}\right) - \Phi\left(\frac{-\frac{s}{2}-0}{\sigma}\right)$

$$= \Phi\left(\frac{s}{2\sigma}\right) - \left(1 - \Phi\left(\frac{s}{2\sigma}\right)\right)$$

$$= 2\Phi\left(\frac{s}{2\sigma}\right) - 1$$

$\therefore \underline{P\_zeros} = P(-\frac{s}{2} < R^* < \frac{s}{2}) = \underline{2\Phi\left(\frac{s}{2\sigma}\right) - 1}$ （频率估计概率）

$$\Rightarrow \quad \Phi\left(\frac{s}{2\sigma}\right) = \frac{1 + P\_zeros}{2}$$

$$\Rightarrow \quad s = 2 \cdot \sigma \cdot \Phi^{-1}\left(\frac{1 + P\_zeros}{2}\right)$$

16. **Amihud measure of illiquidity (*Amihud*).** Following Amihud (2002), the measure is motivated to capture the price impact and is defined as,

$$\text{Amihud} = \frac{1}{N} \sum_{d=1}^{N} \frac{|r_d|}{Q_d},$$

where N is the number of positive-volume days in a given month, r_d the daily return, and Q_d the trading volume on day d, respectively.

17. **An extended Roll's measure (*PI_Roll*).** Goyenko, Holden, and Trzcinka (2009) derive an extended transaction cost proxy measure, which for every transaction cost proxy *tcp* and average daily dollar volume $\overline{Q}$ in the period under observation is defined as

$$\text{PI\_Roll} = \frac{Roll}{\overline{Q}}.$$

18. **An extended FHT measure based on zero returns (*PI_FHT*).**

$$\text{PI\_FHT} = \frac{P\_FHT}{\overline{Q}}.$$

where P_FHT is the modified illiquidity measure based on zero returns (Fong, Holden, and Trzcinka, 2017) and $\overline{Q}$ is the average daily dollar volume in the period under observation.

19. **An extended High-low spread estimator (*PI_HighLow*).**

$$\text{PI\_HighLow} = \frac{P\_HighLow}{\overline{Q}}.$$

where PI_HighLow is the high-low spread estimator following Corwin and Schultz (2012) and $\overline{Q}$ is the average daily dollar volume in the period under observation.

20. **Std.dev of the Amihud measure (Std_Amihud).** The standard deviation of the daily Amihud measure within a month.

21. **Lambda (*PI_Lambda*).** Hasbrouck (2009) proposes Lambda as a high-frequency price impact measure for equities. PI_Lambda (λ) is estimated in the regression,

$$r_\tau = \lambda \cdot \text{sign}(Q_\tau) \cdot \sqrt{|Q_{\tau|}} + \epsilon_\tau,$$

where r_τ is the stock's return and Q_τ is the signed traded dollar volume within the five minute period τ. Following Hasbrouck (2009) and Schestag, Schuster, and Uhrig-Homburg (2016), we take into account the effects of transaction costs on small trades versus large trades (Edwards, Harris, and Piwowar, 2007) and run the adjusted regression,

$$r_i = \alpha \cdot D_i + \lambda \cdot D_i \cdot \sqrt{Q_i} + \epsilon_i,$$

where λ is estimated in the equation above excluding all overnight returns and D_i is an indicator variable of trades defined as the following,

$$D_i = \begin{cases} 1 & \text{if trade } i \text{ is a buy,} \\ 0 & \text{if trade } i \text{ is an interdealer trade,} \\ -1 & \text{if trade } i \text{ is a sell.} \end{cases}$$

22. **Difference of average bid and ask prices (AvgBidAsk).** Following Hong and Warga (2000) and Chakravarty and Sarkar (2003), we use the difference between the average customer buy and the average customer sell price on each day to quantify transaction costs:

$$\text{AvgBidAsk} = \frac{\overline{P_t^{Buy}} - \overline{P_t^{Sell}}}{0.5 \cdot (\overline{P_t^{Buy}} + \overline{P_t^{Sell}})}$$

where $\overline{P_t^{Buy/Sell}}$ is the average price of all customer buy/sell trades on day t. We calculate AvgBidAsk for each day on which there is at least one buy and one sell trade and use the monthly mean as a monthly transaction cost measure.

23. **Interquartile range (TC_IQR).** Han and Zhou (2007) and Pu (2009) use the interquartile range of trade prices as a bid-ask spread estimator. They divide the difference between the 75th percentile P_t^{75th} and the 25th percentile P_t^{25th} of intraday trade prices on day t by the average trade price $\overline{P_t}$ of that day:

$$\text{TC\_IQR} = \frac{P_t^{75th} - P_t^{25th}}{\overline{P_t}},$$

We calculate IQR for each day that has at least three observations and define the monthly measure as the mean of the daily measures.

24. **Round-trip transaction costs (RoundTrip).** Following Feldhütter (2012), we aggregate all trades per bond with the same volumes that occur within a 15-minute time window to a round-trip transaction. We then compute the estimator for round-trip transaction costs as the doubled difference between the lowest and highest trade price for each round-trip transaction. To obtain a relative spread proxy, we divide the round-trip transaction cost estimator by the mean of the maximum and the minimum price. A bond's monthly round-trip measure is then obtained by averaging over all round-trip trades in a month.

25. **Pastor and Stambaugh's liquidity measure (GammaPS, γ_{PS}).** Pástor and Stambaugh (2003) develop a measure for price impact based on price reversals for the equity market. It is given by the estimator for γ in the following regression:

$$r^e_{t+1} = \theta + \psi \cdot r_t + \gamma \cdot sign(r^e_t) \cdot Q_t + \epsilon_t,$$

where r^e_t is the security'e excess return over a market index return, r_t is the security's return and Q_t is the trading volume at day t. For corporate bond market index, we use Merrill Lynch aggregate corproate bond index. γ should be negative and a larger price impact leads to a larger absolute value. As liquidity measures generally assign larger (positive) values to more illiquid bonds, we define $\gamma_{PS} = -\gamma$ expect it to be positively correlated with the other liquidity measures.

26. **Bond market beta (β^{Bond}).** We estimate the bond market beta, β^{Bond}, for each bond from the time-series regressions of individual bond excess returns on the bond market excess returns (MKT^{Bond}) using a 36-month rolling window. We compute the bond market excess return (MKT^{Bond}) as the value-weighted average returns of all corporate bonds in our sample minus the one-month Treasury-bill rate.[17]

27. **Default beta (β^{DEF}).** We estimate the default beta for each bond from the time-series regressions of individual bond excess returns on the bond market excess returns (MKT^{Bond}) and the default factor using a 36-month rolling window. Following Fama and French (1993), the default factor (DEF) is defined as the difference between the return on a market portfolio of long-term corporate bonds (the composite portfolio on the corporate bond module of Ibbotson Associates) and the long-term government bond return.

28. **Term beta (β^{TERM}).** We estimate the default beta for each bond from the time-series regressions of individual bond excess returns on the bond market excess returns (MKT^{Bond}) and the term factor using a 36-month rolling window. Following Fama and French (1993), the term factor (TERM) is defined as the difference between the monthly long-term government bond return (from Ibbotson Associates) and the one-month Treasury bill rate.

29. **Illiquidity beta (β^{LWW}).** Following Lin, Wang, and Wu (2011), it is estimated as the exposure to the bond illiquidity factor, which is defined as the average return difference between the high liquidity beta portfolio (decile 10) and the low liquidity beta portfolio (decile 1).

30. **Downside risk beta (β^{DRF}).** Following Bai, Bali, and Wen (2019), for each bond and each month in our sample, we estimate the factor beta from the monthly rolling regressions of excess bond returns on the downside risk factor (DRF) over a 36-month fixed window after controlling for the bond market factor (MKT^{Bond}).

31. **Credit risk beta (β^{CRF}).** Similar to the construction of downside risk beta, for each bond and each month in our sample, we estimate the factor beta from the monthly rolling regressions of excess bond returns on the credit risk factor (CRF) over a 36-month fixed window after controlling for the bond market factor (MKT^{Bond}).

32. **Illiquidity risk beta (β^{LRF}).** Similar to the construction of downside risk and credit risk beta, for each bond and each month in our sample, we estimate the factor beta from the monthly rolling regressions of excess bond returns on the liquidity risk factor (LRF) over a 36-month fixed window after controlling for the bond market factor (MKT^{Bond}).

33. **Volatility beta (β^{VIX}).** Following Chung, Wang, and Wu (2019), we estimate the following bond-level regression

$$R_{i,t} = \alpha_i + \beta_{1,i}MKT_t + \beta_{2,i}SMB_t + \beta_{3,i}HML_t + \beta_{4,i}DEF_t + \beta_{5,i}TERM_t + \beta_{6,i}\Delta VIX_t + \epsilon_{i,t},$$

where $R_{i,t}$ is the excess return of bond i in month t, and MKT_t, SMB_t, HML_t, DEF_t, $TERM_t$, and ΔVIX_t denote the aggregate corporate bond market, the size factor, the book-to-market factor, the default factor, the term factor, and the market volatility risk factor, respectively.

34. **Macroeconomic Uncertainty Beta (β^{UNC}).** Following Bali, Subrahmanyam, and Wen (2020), for each bond-month in our sample, we estimate the uncertainty beta from monthly rolling regressions of excess bond returns on the change in the economic uncertainty index (ΔUNC) and the excess bond market returns (MKT), using the past 24 to 36 months of data (as available):

$$R_{i,t} = \alpha_{i,t} + \beta_{i,t}^{UNC} \cdot \Delta UNC_t + \beta_{i,t}^{MKT} \cdot MKT_t + \epsilon_{i,t},$$

where $R_{i,t}$ is the excess return of bond i in month t, ΔUNC_t is the change in the economic uncertainty index in month t based on Jurado, Ludvigson, and Ng (2015), MKT_t is the aggregate corporate bond market, $\beta_{i,t}^{UNC}$ is the uncertainty beta of bond i in month t.

35. **Short-term reversal (*REV*).** The bond return in previous month.

36. **Six-month momentum (*MOM6*).** Following Jostova et al. (2013), it is defined as the cumulative bond returns over months from $t-7$ to $t-2$ (formation period), skipping the short-term reversal month.

37. **Twelve-month momentum (*MOM12*).** It is defined as the cumulative bond returns over months from $t-12$ to $t-2$ (formation period), skipping the short-term reversal month.

38. **Long-term reversal (*LTR*).** Following Bali, Subrahmanyam, and Wen (2021), it is defined as the past 36-month cumulative returns from $t-48$ to $t-13$, skipping the 12-month momentum and short-term reversal month.

39. **Volatility (*VOL*).** Following Bai, Bali, and Wen (2016), it is estimated using a 36-month rolling window for each bond in our sample

$$VOL_{i,t} = \frac{1}{n-1}\sum_{t=1}^{n}(R_{i,t} - \overline{R}_i)^2.$$

40. **Skewness (*SKEW*).** Similar to the construction of volatility, skewness is estimated using a 36-month rolling window for each bond in our sample

$$SKEW_{i,t} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{R_{i,t} - \overline{R}_i}{\sigma_{i,t}}\right)^3.$$

41. **Kurtosis (*KURT*).** Similar to the construction of volatility and skewness, kurtosis is estimated using a 36-month rolling window for each bond in our sample

$$KURT_{i,t} = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{R_{i,t} - \overline{R}_i}{\sigma_{i,t}}\right)^4 - 3.$$

42. **Co-skewness (*COSKEW*).** Harvey and Siddique (2000), Mitton and Vorkink (2007), and Boyer, Mitton, and Vorkink (2010) provide empirical support for the three-moment asset pricing models that stocks with high co-skewness, high idiosyncratic skewness, and high expected skewness have low subsequent returns. Following the aforementioned studies, we decompose total skewness into two components; systematic skewness and idiosyncratic skewness, which are estimated based on the following time-series regression for each bond using a 36-month rolling window:

$$R_{i,t} = \alpha_i + \beta_i \cdot R_{m,t} + \gamma_i \cdot R^2_{m,t} + \varepsilon_{i,t}.$$

where $R_{i,t}$ is the excess return on bond i, $R_{m,t}$ is the excess return on the bond market portfolio, γ_i is the systematic skewness (co-skewness) of bond i.

43. **Idiosyncratic skewness (*ISKEW*).** The idiosyncratic skewness (*ISKEW*) of bond i is defined as the skewness of the residuals ($\varepsilon_{i,t}$) in co-skewness regression equation.