Predicting Corporate Bond Returns: Merton Meets Machine Learning

Turan G. Bali, Amit Goyal, Dashan Huang, Fuwei Jiang, Quan Wen

Working paper,2022

解读者：屠雪永

2023.02.15

Outline

- Introduction
- Methodology
  - Data and variable definitions
- Empirical study
  - Predicting Bond Returns without Hedge Ratios
  - Predicting Bond Returns with Hedge Ratios
- Conclusion

1. Introduction-- Motivation

- The proliferation(激增) of stock characteristics or factors to explain the cross-section of stock returns
  - many stock characteristics
  - data snooping、p-hacking
  - robust forecasting
- Far fewer studies are devoted to predict future returns on corporate bonds
  - default and term betas (Fama and French, 1993; Gebhardt, Hvidkjaer, and Swaminathan, 2005),
  - liquidity risk (Lin, Wang, and Wu,2011), downside risk (Bai, Bali,and Wen, 2019),
  - bond momentum (Jostova, Nikolova, Philipov, and Stahel, 2013), and long-term reversal (Bali, Subrahmanyam, and Wen, 2021a),

1. Introduction-- Motivation

- Using the OLS regressions, whether well-known **equity market anomalies** impact the cross-section of corporate bond returns and find **mixed evidence** on the predictability
    - bondholders are more sensitive to downside risk
    - the high correlation between many of the stock and bond characteristics

- Previous studies, in general, rely on the **reduced-form** approach that examines cross-sectional bond return predictability, **without explicitly linking** the functional forms of bond and stock expected returns.
    - the same underlying assets of the firm
    - Merton (1974) explains how bonds and stocks should be jointly priced

1. Introduction-- Questions

- Do machine learning models substantially improve the out-of-sample performance of bond characteristics in predicting bond returns?

Yes

- Do Stock Characteristics predict bond returns?

Yes

- Do Stock Characteristics Improve the Predictive Power of Bond Characteristics for Future Bond Returns without restrictions?

No

- Is there a significant **improvement** in the performance of machine learning models when imposing an **economic structure** from the Merton model?

Yes

1. Introduction-- Contribution

- We provide a comprehensive study on the cross-sectional predictability of corporate bond returns using a large set of stock and bond characteristics.

- We first build a comprehensive data library of 43 corporate bond-level characteristics.

- We investigate the predictability of bond returns using **Merton (1974) model** with hedge ratios

2. Methodology

➢ 2.1 Theoretical Motivation

- Model to analyze the **stock-bond connection**
  - **Merton (1974)** structural credit risk model,
  - its extension in **Du, Elkamhi, and Ericsson (2019)**

- Assume that the value of the assets of the firm, V_t, is governed by the following stochastic processes

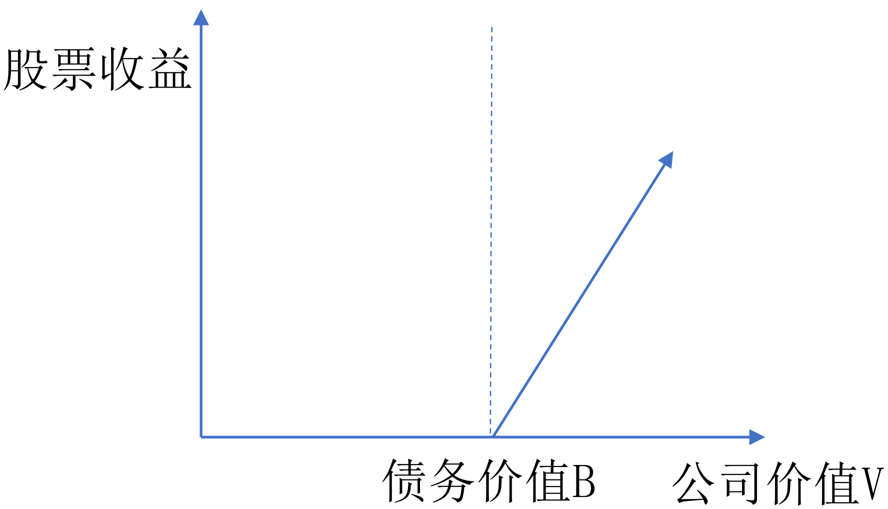
$$\begin{aligned} dV_t &= rV_t dt + \sigma_t V_t dW_t \\ d\sigma_t^2 &= \kappa(\theta - \sigma_t^2)dt + \gamma\sigma_t dZ_t, \end{aligned}$$

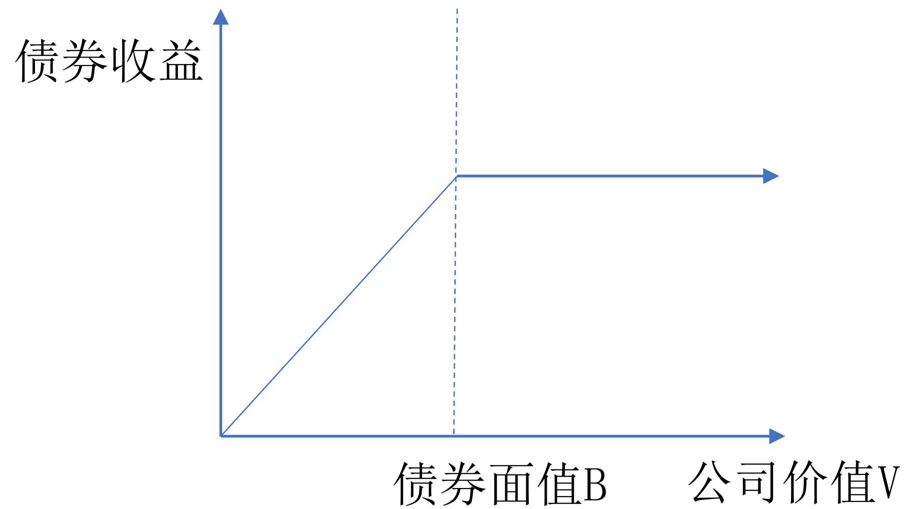
  - r is the risk-free rate, processes $\{W_t\}$ and $\{Z_t\}$ are two standard Brownian motions, κ is the speed of mean reversion, θ is the long-run mean variance, and γ is the volatility parameter for asset variance.

2. Methodology

➢ 2.1 Theoretical Motivation

- Stock: a long position in call option(看涨期权的多头)
- Bond: a short position in a put option(看跌期权的空头)





- The equity prices

$$E_t = V_t N(d_1) - Be^{-r(T-t)} N(d_2) - \frac{\sqrt{\theta}\gamma}{8\kappa} \cdot \eta_t,$$

- Bond prices

$$D_t = V_t - E_t$$

2. Methodology

$$E_t = V_t N(d_1) - Be^{-r(T-t)} N(d_2) - \frac{\sqrt{\theta}\gamma}{8\kappa} \cdot \eta_t,$$

$$D_t = V_t - E_t$$

➢ 2.1 Theoretical Motivation

- The hedge ratio (Schaefer and Strebulaev, 2008)

$$\begin{aligned} h_t &\stackrel{\text{def}}{=} \frac{\partial D_t / \partial V_t}{\partial E_t / \partial V_t} \times \frac{E_t}{D_t} \\ &= \frac{1 - N(d_1) + \gamma^2 \zeta_t}{N(d_1) - \gamma^2 \zeta_t} \frac{E_t}{D_t}. \end{aligned}$$

- The equity and bond returns have the following relationship

$$\frac{dD_t}{D_t} - h_t \frac{dE_t}{E_t} = \alpha_t dt.$$

- Prediction of bond returns involves three components
  - (i) predicting the hedge ratio,
  - (ii) predicting the stock return,
  - (iii) predicting the ‘residual’ bond return.

2. Methodology

$$\frac{dD_t}{D_t} - h_t \frac{dE_t}{E_t} = \alpha_t dt.$$

- 2.2 Prediction Framework

- The excess return of asset i at time $t + 1$

$$R_{it+1} = E_t(R_{it+1}) + e_{it+1},$$

- let RB and RS denote the realized bond and stock return

$$\begin{aligned} RB_{it+1} &= E_t[RB_{it+1}] + eB_{it+1} \\ RS_{it+1} &= E_t[RS_{it+1}] + eS_{it+1}, \end{aligned}$$

where eB and eS are the unexpected bond return and stock return

$$E_t(RB_{it+1}) = h_{it} \times E_t(RS_{it+1}) + \alpha_{it}.$$

$$\begin{aligned} RB_{it+1} &\equiv E_t(RB_{it+1}) + eB_{it+1} \\ &= h_{it} \times E_t(RS_{it+1}) + \alpha_{it} + eB_{it+1} \\ &= h_{it} \times RS_{it+1} + \alpha_{it} + (eB_{it+1} - h_{it} \times eS_{it+1}). \end{aligned}$$

2. Methodology

$$\begin{aligned} RB_{it+1} &\equiv E_t(RB_{it+1}) + eB_{it+1} \\ &= h_{it} \times E_t(RS_{it+1}) + \alpha_{it} + eB_{it+1} \\ &= h_{it} \times RS_{it+1} + \alpha_{it} + (eB_{it+1} - h_{it} \times eS_{it+1}). \end{aligned}$$

➢ 2.2 Prediction Framework

- Define $\boldsymbol{RBmRS_{it+1}}$ as the difference between realized bond return (RB) and the product of the hedge ratio and realized stock return ($h \times RS$):

$$\begin{aligned} RBmRS_{it+1} &\stackrel{\text{def}}{=} RB_{it+1} - h_{it} \times RS_{it+1} \\ &= \alpha_{it} + (eB_{it+1} - h_{it} \times eS_{it+1}). \end{aligned}$$

- Taking expectation, we see that $E_t(RBmRS_{it+1}) = \alpha_{it}$

$$E_t(RB_{it+1}) = h_{it} \times E_t(RS_{it+1}) + E_t(RBmRS_{it+1})$$

$$E_t(R_{it+1}) = \phi(X_{it})$$

$\phi(\cdot)$ is a flexible function of asset i's P-dimensional characteristics

2. Methodology

$$E_t(RB_{it+1}) \quad = \quad h_{it} \times E_t(RS_{it+1}) + E_t(RBmRS_{it+1})$$

➢ 2.2 Prediction Framework

- Three variations of predicting bond returns:
- 1. Without the hedge ratios

$$E_t(RB_{it+1}) = f_1(X_{it}),$$

  - where the characteristics X include combinations of bond characteristics, XB, and stock characteristics, XS
- 2.With regression-based hedge ratios

$$E_t(RB_{it+1}) = f_2(X_{it}) = \hat{h}_{it} \times \psi_1(X_{it}) + \psi_2(X_{it})$$

  - estimate hedge ratios via regressions of bond returns on stock returns
- 3.With machine learning-based hedge ratios

$$E_t(RB_{it+1}) = f_3(X_{it}) = \phi_3(X_{it}) \times \phi_1(X_{it}) + \phi_2(X_{it}).$$

  - let the hedge ratio itself be a function of characteristics.

2. Methodology

- 2.3 Machine Learning and Performance Evaluation
  - The ordinary least squares (OLS)
  - **Penalized linear regression** methods such as LASSO, ridge regression (Ridge), and elastic net (ENet)
  - **Dimension reduction** techniques such as principal component analysis (PCA) and partial least square (PLS)
  - Random forests (RF);
  - Feed-forward neural network (FFN)
  - Long short-term memory neural network (LSTM)

2. Methodology

➢ 2.3 Machine Learning and Performance Evaluation

- Assess the predictive power of individual bond return predictors

$$R_{OS}^2 = 1 - \frac{\sum_{(i,t)\in\mathcal{T}_3}\left(r_{it+1} - \hat{r}_{it+1}\right)^2}{\sum_{(i,t)\in\mathcal{T}_3} r_{it+1}^2}$$

- To compare the out-of-sample predictive power of two methods, the modified Diebold and Mariano (1995) test

$$\mathrm{DM}_{12} = \bar{d}_{12}/\hat{\sigma}_{\bar{d}},$$

$$d_{12,t+1} = \frac{1}{n_{3,t+1}} \sum_{i=1}^{n_3} \left(\left(\hat{e}_{it+1}^{(1)}\right)^2 - \left(\hat{e}_{it+1}^{(2)}\right)^2 \right)$$

$d_{12,t+1}$ is the forecast error differential between the two methods.

3. Data and Variable Definitions

- Data: TRACE(交易信息), Mergent FISD(基本信息), 200206-201712
- The filtering criteria proposed by Bai, Bali, and Wen (2019):
  - are not listed or traded in the US public market;
  - structured notes, mortgage backed/asset backed/agency backed/equity-linked;
  - are convertible;
  - trade under $5;
  - have floating coupon rates;
  - have less than one year to maturity.
  - are canceled
  - have more than a two-day settlement
  - have a trading volume smaller than $10,000

3. Data and Variable Definitions

➢ 3.1 Corporate Bond Return

- The monthly corporate bond return at time t

$$r_{it} = \frac{P_{it} + AI_{it} + C_{it}}{P_{it-1} + AI_{it-1}} - 1,$$

  - where P_{it} is the transaction price, AI_{it} is accrued interest, and C_{it} is the coupon payment

- bond i's excess return

$$R_{it} = r_{it} - r_{ft}$$

  - where r_{ft} is the risk-free rate proxied by the one-month Treasury bill rate.

3. Data and Variable Definitions

- 3.2 Corporate Bond and Equity Characteristics
- 94 stock-level predictors
- 43 corporate bond characteristics
  - (i) bond-level characteristics such as issuance size, age, credit rating, time-to-maturity, and duration,
  - (ii)proxies of corporate bond downside risk,
  - (iii) proxies of bond-level illiquidity and liquidity risk,
  - (iv) proxies of systematic risk such as default and term betas and volatility betas,
  - (v) past bond return characteristics such as bond momentum, short-term reversal, and long-term reversal,
  - (vi) distributional characteristics including return volatility, skewness, and kurtosis.

4. Predicting Bond Returns without Hedge Ratios

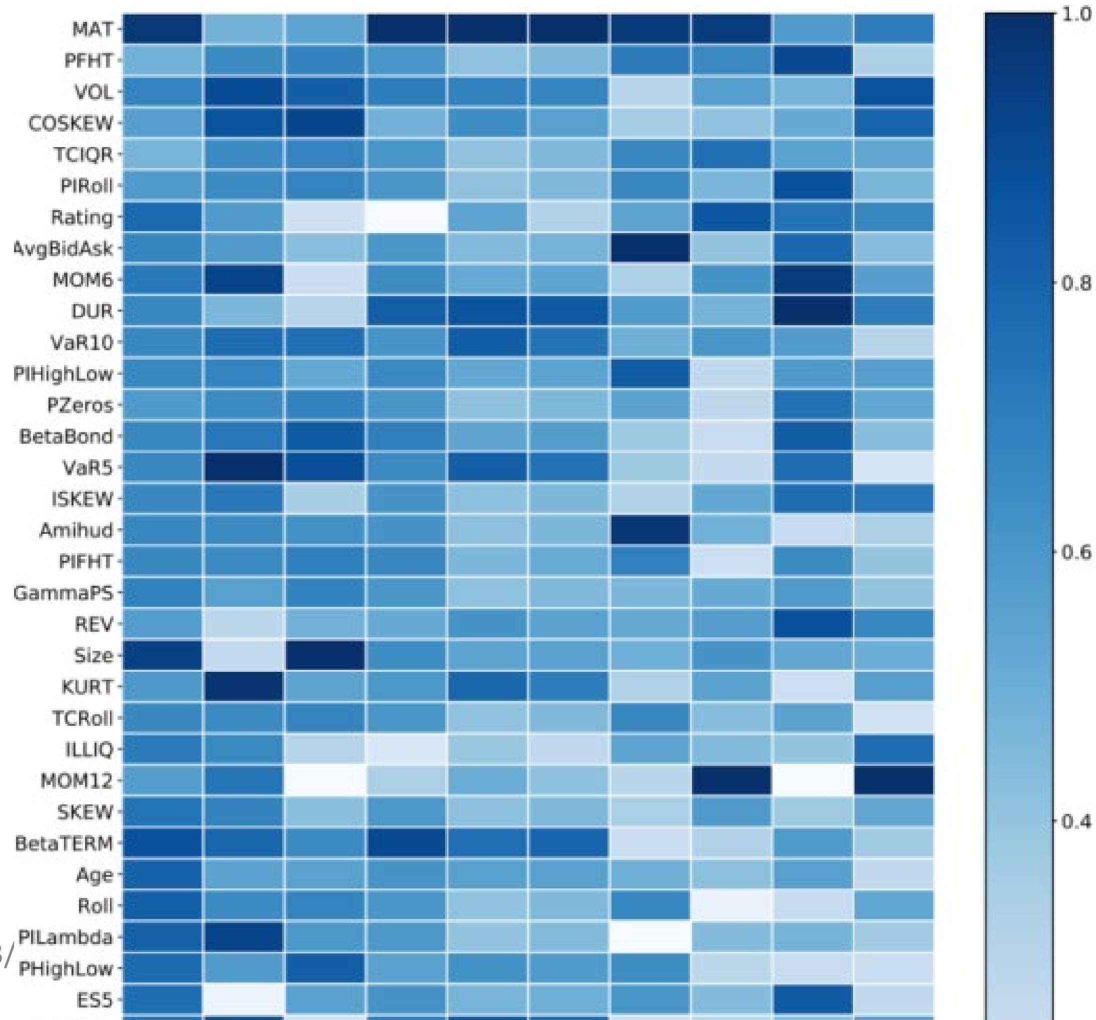
- 4.1 Using only Bond Characteristics
  - 4.1.1 Out-of-Sample Predictive Power

	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
					Panel A: Out-of-sample R^2_{OS}					
R^2_{OS}	−3.36	2.07	2.03	1.85	1.89	1.87	2.19	2.37	2.28	2.09
					Panel B: Comparison of monthly out-of-sample prediction using Diebold-Mariano tests					
OLS		**3.07**	**2.89**	**3.45**	**3.53**	**3.59**	**3.82**	**3.85**	**3.28**	**3.38**
PCA			1.14	−1.32	−1.26	−1.40	**2.10**	1.78	0.28	1.85
PLS				−0.79	−0.57	−0.65	1.78	1.70	0.13	1.14
LASSO					0.44	0.40	1.60	1.18	0.86	**2.05**
Ridge						0.15	1.78	1.96	0.86	**2.00**
Enet							1.81	1.08	0.86	**2.10**
RF								1.10	1.74	1.91
FFN									−0.80	1.20
LSTM										1.15

4. Predicting Bond Returns without Hedge Ratios

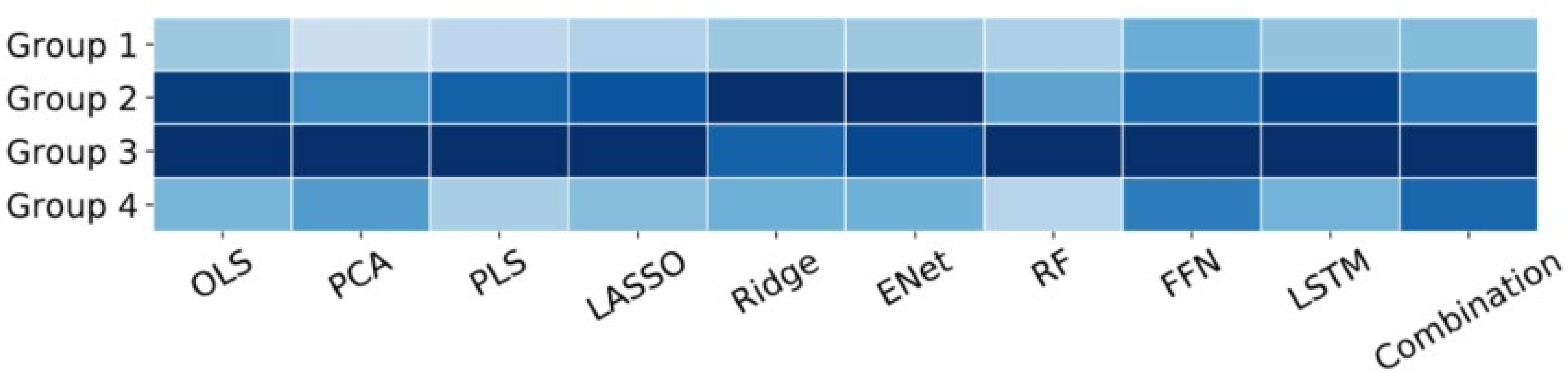
- 4.1 Using only Bond Characteristics
  - 4.1.2 Which Bond Characteristics Matter?

•



4. Predicting Bond Returns without Hedge Ratios

- 4.1 Using only Bond Characteristics



(i) bond characteristics related to **interest rate risk** such as duration (DUR) and time-to-maturity (MAT),

(ii) **risk measures** such as **downside risk** proxied by Value-at-Risk (VaR) and expected shortfall (ES), total return volatility (VOL), and **systematic risk** related to bond market beta, default beta, term beta, and economic uncertainty beta,

(iii) bond-level **illiquidity** measures such as the average bid and ask price (AvgBidAsk), Amihud and Roll's measures of illiquidity,

(iv) **past return characteristics** related to bond momentum (MOM), short-term reversal (STR), and long-term reversal (LTR)

4. Predicting Bond Returns without Hedge Ratios

- 4.1 Using only Bond Characteristics
  - 4.1.3 Machine Learning based Long-Short Portfolios

	Low	2	3	4	5	6	7	8	9	High	High−Low	
OLS	0.60	0.67	0.62	0.61	0.56	0.59	0.63	0.60	0.53	0.76	0.16	(1.38)
PCA	0.68	0.61	0.68	0.64	0.65	0.70	0.63	0.67	0.65	1.19	0.51	(2.51)
PLS	0.51	0.55	0.57	0.55	0.57	0.58	0.67	0.65	0.68	1.14	0.63	(2.86)
LASSO	0.57	0.50	0.47	0.40	0.42	0.42	0.46	0.59	0.58	0.96	0.39	(2.54)
Ridge	0.58	0.53	0.46	0.46	0.52	0.45	0.62	0.60	0.67	0.91	0.33	(2.15)
Enet	0.54	0.52	0.48	0.35	0.43	0.41	0.45	0.58	0.55	0.97	0.43	(2.67)
RF	0.57	0.69	0.54	0.51	0.52	0.50	0.59	0.55	0.49	1.37	0.79	(2.78)
FFN	0.61	0.63	0.48	0.55	0.49	0.59	0.50	0.59	0.56	1.36	0.75	(2.61)
LSTM	0.53	0.64	0.60	0.53	0.47	0.55	0.56	0.62	0.58	1.32	0.79	(3.33)
Combination	0.71	0.63	0.58	0.50	0.52	0.60	0.65	0.61	0.59	1.38	0.67	(3.41)

4. Predicting Bond Returns without Hedge Ratios

- 4.1 Using only stock characteristics

	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
	Panel A: R^2_{OS} using stock characteristics									
Using $f_1(XS)$	−3.09	1.70	1.71	1.61	1.57	1.62	1.80	1.88	2.00	2.02
	Panel B: Performance of machine learning High−Low bond portfolio using stock characteristics									
Using $f_1(XS)$	0.02	0.36	0.43	0.24	0.26	0.24	0.43	0.48	0.52	0.52
	(0.12)	(2.35)	(2.67)	(2.12)	(2.11)	(2.03)	(2.28)	(2.25)	(3.09)	(3.13)

4. Predicting Bond Returns without Hedge Ratios

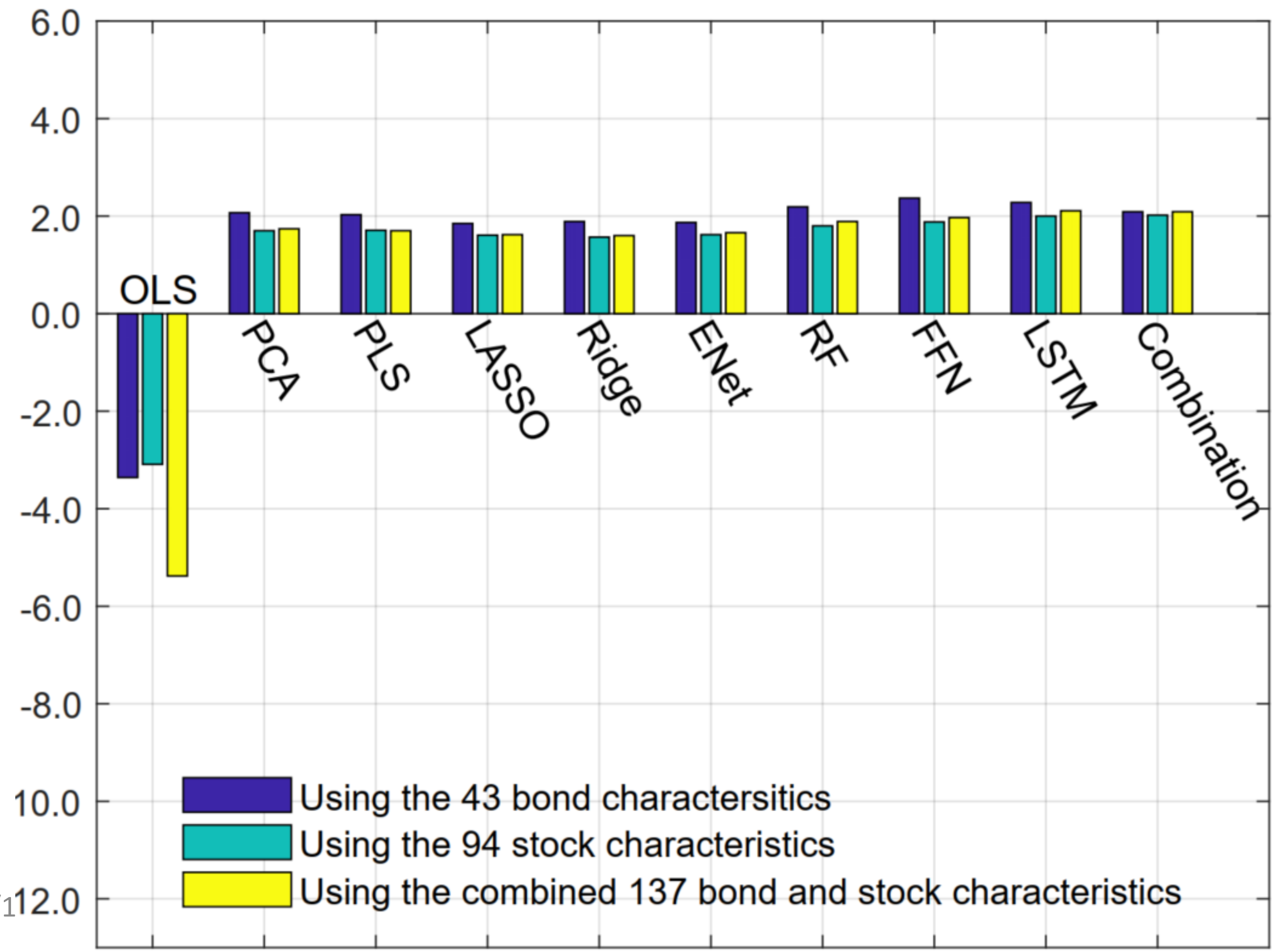
➢ 4.3 Do Stock Characteristics Improve the Predictive Power of Bond Characteristics for Future Bond Returns?

	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
	Panel A: R^2_{OS} using stock and bond characteristics									
Using $f_1(XB, XS)$	−5.38	1.74	1.70	1.62	1.60	1.66	1.89	1.97	2.11	2.09
	Panel B: Comparing machine learning High−Low bond portfolio									
Using $f_1(XB, XS)$	0.11	0.51	0.57	0.41	0.37	0.44	0.68	0.64	0.71	0.65
	(1.18)	(2.45)	(2.35)	(2.15)	(2.13)	(2.25)	(3.13)	(3.11)	(3.19)	(3.08)
Using $f_1(XB, XS)$ − Using $f_1(XB)$	−0.05	0.00	−0.06	0.02	0.04	0.01	−0.11	−0.11	−0.08	−0.02
	(−0.97)	(0.02)	(−1.01)	(0.22)	(0.99)	(0.68)	(−1.26)	(−1.35)	(−1.45)	(−0.92)
Using $f_1(XB, XS)$ − Using $f_1(XS)$	0.09	0.15	0.14	0.18	0.11	0.21	0.25	0.16	0.19	0.13
	(2.33)	(1.81)	(1.88)	(1.78)	(1.38)	(1.77)	(2.86)	(2.22)	(2.00)	(2.15)

4. Predicting Bond Returns without Hedge Ratios

- 4.3 Do Stock Characteristics Improve the Predictive Power of Bond Characteristics for Future Bond Returns?

- R_{OS}^2



4. Predicting Bond Returns without Hedge Ratios

- 4.4 Robustness Checks

  - Transaction Cost
  - Time-varying Performance
  - Maturity-matched Bond Excess Returns
  - Removing Financial Firms

5. Predicting Bond Returns with Regression-Based Hedge Ratios

- 1.Estimate hedge ratios $\hat{h}_{i,t}$ based on the 36-month rolling window regression

$$RB_{is} = \alpha_i + \boxed{h_{it}} RS_{is} + eB_{is}, \quad s = t-35, \ldots, t,$$

- 2. Calculate residual

$$\boxed{RBmRS_{it+1}} = RB_{it+1} - \hat{h}_{it} \times RS_{it+1}$$

- 3. Run separate machine learning models to predict

$$\begin{aligned} E_t(RBmRS_{it+1}) &= \psi_2(XB_{it}) \\ E_t(RS_{it+1}) &= \psi_1(XS_{it}). \end{aligned}$$

- 4. Prediction for expected bond return

$$E_t(RB_{it+1}) = f_2(XB_{it}, XS_{it}, \hat{h}_{it}) = \hat{h}_{it} \times \psi_1(XS_{it}) + \psi_2(XB_{it}).$$

5. Predicting Bond Returns with Regression-Based Hedge Ratios

	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
	Panel A: R^2_{OS}									
Using $f_2(XB, XS, \hat{h})$	−4.37	2.28	2.88	1.93	1.95	1.95	3.05	3.11	4.89	4.95
	Panel B: Comparison of monthly out-of-sample prediction using Diebold-Mariano tests									
Using $f_2(XB, XS, \hat{h})$ − Using $f_1(XB)$	−1.01	**0.21**	**0.85**	0.08	0.06	0.08	**0.86**	**0.74**	**2.61**	**2.86**
Using $f_2(XB, XS, \hat{h})$ − Using $f_1(XB, XS)$	1.01	**0.54**	**1.18**	**0.31**	**0.35**	**0.29**	**1.16**	**1.14**	**2.78**	**2.86**

High-Low return	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
Using $f_2(XB, XS, \hat{h})$	0.18	**0.64**	**0.69**	**0.55**	**0.57**	**0.57**	**0.86**	**0.89**	**0.92**	**0.84**
	(1.07)	(2.16)	(2.33)	(2.60)	(2.75)	(2.77)	(3.01)	(2.68)	(2.69)	(3.27)
Using $f_2(XB, XS, \hat{h})$	0.02	**0.13**	0.06	**0.16**	**0.24**	**0.14**	**0.07**	**0.14**	**0.13**	**0.17**
− Using $f_1(XB)$	(0.35)	(2.35)	(1.93)	(2.27)	(2.45)	(2.36)	(2.04)	(2.54)	(2.38)	(2.81)
Using $f_2(XB, XS, \hat{h})$	0.07	**0.13**	**0.12**	**0.14**	**0.20**	**0.13**	**0.18**	**0.25**	**0.21**	**0.19**
− Using $f_1(XB, XS)$	(0.76)	(2.44)	(2.87)	(2.15)	(2.43)	(2.22)	(2.36)	(2.77)	(2.63)	(2.60)

6. Predicting Bond Returns with Machine Learning Based Hedge Ratios

$$RB_{is} = \alpha_i + \boxed{h_{it}} RS_{is} + eB_{is}, \quad s = t-35, \ldots, t,$$

- 1.Estimate hedge ratios $\hat{h}(XB_{it})$ based on the 36-month rolling window regression

$$RB_{is} = h(XB_{is-1})RS_{is} + uB_{is}, \quad s = t-35, \ldots, t,$$

[16] As an illustrative example, consider prediction using neural network. Given XB_{is-1}, the machine generates K units of neurons for each layer l as $XB_l^{(k)} = g\left(\theta_l^{(k)} XB_{is-1}\right)$, where $g(\cdot)$ is the nonlinear activation function. Then, in the last layer, we multiply each $XB_L^{(k)}$ by RS_{is}. The output is the fitted value $\widehat{RB}_{is} = g\left(XB_L^{(k)} RS_{is}\right) = \hat{h}(XB_{is-1}) \times RS_{is}$. We can calculate the out-of-sample fitted value as $\hat{E}(RB_{it+1}|XB_{it}, RS_{it+1}) = \hat{h}(XB_{it}) \times RS_{it+1}$. When needed, the hedge ratio itself can be recovered by 'setting' the stock return to be one to obtain $\hat{h}(XB_{it}) = \hat{E}(RB_{it+1}|XB_{it}, 1)$.

6. Predicting Bond Returns with Machine Learning Based Hedge Ratios

•

	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
	Panel A: R^2_{OS}									
Using $f_3(XB, XS, \hat{h}(XB))$	−4.59	2.35	3.07	2.04	2.05	2.05	3.30	3.53	5.67	5.70
	Panel B: Comparison of out-of-sample prediction using Diebold-Mariano tests									
Using $f_3(XB, XS, \hat{h}(XB))$ − Using $f_1(XB)$	−1.23	**0.28**	**1.04**	**0.19**	**0.16**	**0.18**	**1.11**	**1.14**	**3.39**	**3.61**
Using $f_3(XB, XS, \hat{h}(XB))$ − Using $f_1(XB, XS)$	0.79	**0.61**	**1.37**	**0.42**	**0.45**	**0.39**	**1.41**	**1.56**	**3.56**	**3.61**
Using $f_3(XB, XS, \hat{h}(XB))$ − Using $f_2(XB, XS, \hat{h})$	−0.22	**0.07**	**0.19**	**0.11**	**0.10**	**0.10**	**0.25**	**0.42**	**0.78**	**0.75**

High-Low return	(1) OLS	(2) PCA	(3) PLS	(4) LASSO	(5) Ridge	(6) ENet	(7) RF	(8) FFN	(9) LSTM	(10) Combination
Using $f_3(XB, XS, \hat{h}(XB))$	0.16	**0.65**	**0.71**	**0.54**	**0.57**	**0.58**	**0.89**	**0.93**	**1.00**	**0.89**
	(0.53)	(2.61)	(2.49)	(2.49)	(2.52)	(2.47)	(2.71)	(2.84)	(3.22)	(4.68)
Using $f_3(XB, XS, \hat{h}(XB))$	0.00	**0.14**	0.08	**0.15**	**0.24**	**0.15**	**0.10**	**0.18**	**0.21**	**0.22**
− Using $f_1(XB)$	(0.02)	(2.14)	(1.81)	(2.43)	(2.55)	(2.61)	(2.05)	(2.07)	(2.12)	(2.41)
Using $f_3(XB, XS, \hat{h}(XB))$	0.05	**0.14**	**0.14**	**0.13**	**0.20**	**0.14**	**0.21**	**0.33**	**0.29**	**0.24**
− Using $f_1(XB, XS)$	(0.76)	(2.25)	(2.41)	(2.42)	(2.56)	(2.41)	(2.21)	(2.44)	(2.51)	(2.75)
Using $f_3(XB, XS, \hat{h}(XB))$	−0.02	0.01	0.02	−0.01	0.00	0.01	0.03	0.04	0.08	0.05
− Using $f_2(XB, XS, \hat{h})$	(−0.22)	(0.21)	(0.44)	(−0.10)	(0.02)	(0.10)	(0.43)	(0.54)	(1.15)	(0.79)

6. Predicting Bond Returns with Machine Learning Based Hedge Ratios

- Comparison of hedge ratios

	Regression-based	Machine learning-based									
		OLS	PCA	PLS	LASSO	Ridge	Enet	RF	FFN	LSTM	Combination
All	0.051	0.097	0.053	0.054	0.053	0.053	0.054	0.055	0.057	0.056	0.054
Investment-grade (Rating $\leq$ 10)	0.033	0.064	0.032	0.031	0.033	0.032	0.032	0.031	0.031	0.031	0.031
Non-investment-grade (Rating $>$ 10)	0.018	0.033	0.021	0.023	0.021	0.021	0.021	0.024	0.025	0.025	0.023
Short-maturity (1 $\leq$ Maturity $<$ 3)	0.003	0.007	0.003	0.003	0.004	0.004	0.004	0.004	0.005	0.004	0.004
Medium-maturity (3 $\leq$ Maturity $<$ 7)	0.024	0.049	0.026	0.027	0.026	0.026	0.026	0.026	0.027	0.026	0.026
Long-maturity (Maturity $\geq$ 7)	0.024	0.042	0.024	0.024	0.023	0.023	0.024	0.025	0.025	0.025	0.024

V. Conclusion

- **The machine learning methods substantially improve the out-of-sample performance in predicting future bond returns**
- Using the **reduced-form** approach, the incremental **improvement** of stock characteristics relative to bond characteristics is **economically and statistically small** in forecasting future bond returns.
- Our work highlights the importance of explicitly imposing the **dependence between expected bond and stock returns** when investigating expected bond returns