Benchmarking Individual Corporate Bonds *

Xin He † Guanhao Feng ‡ Junbo Wang § Chunchi Wu ¶

February 1, 2021

Abstract

We propose an alternative approach to the linear factor model to estimate and decompose individual corporate bond risk premia. Using characteristic-based benchmarks to evaluate assets' performance is a common practice in empirical asset pricing (Daniel et al., 1997), while factor models are difficult to implement for individual bonds with limited observations. To solve the high-dimensional sort difficulty proposed in Cochrane (2011), we introduce a *benchmark combination model (BCM)* that combines multiple basis portfolios to price bonds. With a non-arbitrage objective, our approach minimizes cross-sectional pricing errors and acts as a linear pricing kernel as the factor model. There are three major empirical findings. First, our BCM model outperforms factor models in pricing individual corporate bond returns for multiple performance measures. Second, we find Credit Rating, Downside Risk, and Short-Term Reversal are three primary sources of the individual bond risk premia. Finally, we incorporate machine learning forecasts into the BCM model, which delivers strong return prediction performance.

Key Words: Characteristic-based Benchmark, High-Dimensional Sort, Corporate Bond, Risk Premia, Forecast Combination, Machine Learning, Return Predictability.

*We appreciate insightful comments from Li Guo, Xiaohu Wang, and Zhonghao Fu. We are also grateful for helpful comments from seminar and conference participants at Fudan.

†College of Business, City University of Hong Kong. E-mail address: xin.he@my.cityu.edu.hk.

‡College of Business, City University of Hong Kong. E-mail address: gavin.feng@cityu.edu.hk.

§College of Business, City University of Hong Kong. E-mail address: jwang2@cityu.edu.hk.

¶School of Management, State University of New York at Buffalo. E-mail address: chunchiw@buffalo.edu.

1 Introduction

Asset pricing models study why different assets earn different expected returns. When the market is efficient, expected excess returns (for equities, bonds, and commodities) are supposed to be entirely explained by different risk exposures, and consequently, pricing errors are zero. With the implication of this non-arbitrage condition, asset pricing models are useful to set up benchmark expected returns and evaluate if an asset is underpriced or overpriced over the benchmark.

On the one hand, when evaluating a corporate bond's performance, an investor references the same rating portfolio performance. One might classify individual corporate bonds into five consecutive credit rating buckets (AAA/AA/A/BBB/Junk). All bonds in the same benchmark are believed to bear a similar risk, and their benchmark-adjusted performances show the fairness of asset pricing. In addition to the credit ratings, investors also consider other bond characteristics, such as duration and downside risk. These benchmark evaluations are well established to practitioners and academia and discussed in Section 1.1. If one wants to decompose asset risk premia of excess returns further, the benchmark evaluation becomes inconvenient to handle multiple characteristics.

On the other hand, another typical return benchmark construction is the parametric risk factor model. Factor models have also become the standard tool to estimate and decompose asset risk premia in asset pricing. Besides the superior statistical properties, the non-arbitrage restriction can be formulated quantitatively under the factor model. When one only uses tradable factors to build the model, the intercept alpha becomes the pricing error. According to ICAPM of Merton (1973), a combination of common factors captures the cross-section of expected returns, and the regression intercept should be zero. However, one drawback of the factor model is its regression estimation accuracy relies on the sample size. Therefore, most empirical finance studies focus on estimating factor models for portfolios instead of individual asset returns.

In this paper, we investigate this problem and provide an alternative approach to price individual corporate bonds. The goal of our paper is threefold. First, we tackle the high-dimensional sort difficulty in the cross-section of individual corporate bond returns. Second, we evaluate the model fitness with an economic objective, pricing errors, to compare our approach with existing factor models. Finally, besides the pricing performance, we demonstrate the prediction performance for predicting individual corporate bond returns.

1.1 Sorting Mechanism and Challenge

Sorting corporate bonds into consecutive buckets by their characteristics, such as credit ratings, is a standard benchmark evaluation for individual corporate bonds. Benchmark evaluation via sorting is a non-parametric approach and also well established in the equity market. Daniel et al. (1997) apply characteristic-sorted benchmark portfolios to evaluate mutual fund performance and price equity returns.[1] The idea is to reduce the vast asset universe's dimension into 125 ($5 \times 5 \times 5$) buckets by market equity, book-to-market ratio, and return momentum. To assess an asset's realized and expected returns, one chooses the benchmark as the corresponding bucket portfolio out of 125. We use Figure 1 to demonstrate the sorting mechanism for the benchmark evaluation.

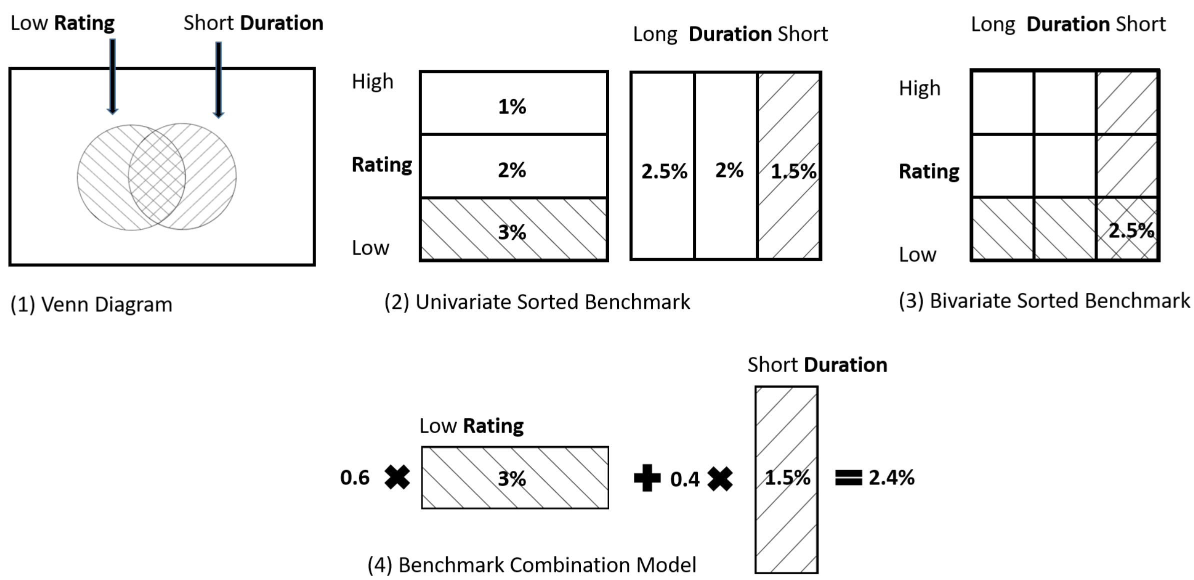

Figure 1: Sorting in Corporate Bonds

This figure demonstrates the sorting mechanism as well as the bench combination model. The numbers on the shaded area denote the values of benchmark returns, for demonstration.

The Venn diagram in sub-figure (1) represents low-rating bonds, short-duration bonds, and the whole bond universe. The intersection is the set of bonds with low ratings and short durations. The univariate sorting in sub-figure (2) demonstrates how one sorts corporate bonds on the credit

[1]The benchmark evaluation for practitioners (i.e., Morningstar) is called "style box analysis," which provides a graphical representation of investing categories for fixed-income or equity investments. The commonly used category pairs are size-value for equities and rating-duration for corporate bonds.

ratings or duration. For example, in the left box of sub-figure (2), the whole bond universe is sorted into three buckets based on the credit ratings, from high to low. The shaded bucket represents the low-rating bonds, equivalent to the left circle in sub-figure (1). One usually includes all bonds from the same bucket in a basis portfolio. For example, the low-rating bonds in the same basis portfolio are believed to have similar default risk exposure, thus sharing similar default risk premia. The bivariate sorting in sub-figure (3) independently displays two boxes in sub-figure (2) into a 3-by-3 style box. The bottom-right bucket is the intersection of low-rating and short-duration bonds, equivalent to the intersection set in sub-figure (1). The portfolio constructed by the bottom-right bucket serves as a benchmark for low-rating and short-duration bonds.

Though the multi-dimensional sort reflects different risk exposures, researchers usually perform the bivariate sort due to the shrinking portfolio size in interactions. The interaction between characteristics in a high-dimensional sort can reduce the observation number in some buckets dramatically. As shown in Table 1, we perform the bivariate-sort of corporate bonds on duration-rating in the upper panel and duration-downside risk in the lower panel. We find a few buckets with less than ten observations, which means the idiosyncratic risk cannot be well-diversified, and the properties of these benchmarks are relatively unstable. Implementing a high-dimensional sort is impossible as questioned in Cochrane (2011). Though the benchmark evaluation via sorting is straightforward, it is infeasible for the risk premia decomposition.

1.2 Our Solution

This paper constructs the benchmark evaluation through a linear combination of basis portfolios (univariate-sorted portfolios), an alternative to the multidimensional-sorted benchmark portfolios. As the sub-figure (4) in Figure 1 shows, we use a linear combination of two basis portfolio returns to replace the corresponding bucket portfolio in a bivariate-sorted "style box."

This linear combination of basis portfolios is one solution to the multidimensional-sort challenge. First, the pricing kernel is still a tradable portfolio because it is a portfolio of sorted portfolios. Second, it does not suffer from the limited sample size for the multidimensional sort. Finally, it reflects different risk exposures using the combination weights. A recent paper of Kelly et al. (2019) suggests sorting on different characteristics provides bootstrap or ensemble samples to dissect the

Table 1: The Number of Observations by Buckets for Bivariate Sorts

This table reports the monthly average number of observations in each bucket for independently bivariate sorts from year 2015 to 2017. We sort individual corporate bonds by Credit Ratings, Duration, and Downside Risk into five buckets correspondingly. The top table shows the number of observations for Rating and Duration bivariate sorts, and the bottom table shows the case for the Duration and Downside Risk.

		Duration					
		1	2	3	4	5	All
Rating	AAA	42	39	32	16	30	158
	AA	38	30	19	15	23	125
	A	114	104	81	92	104	494
	BBB	78	86	106	141	122	532
	Junk	22	34	55	31	18	161
	All	293	293	293	295	296	1471
Downside Risk	1	9	15	69	48	156	296
	2	15	30	71	74	105	295
	3	42	59	79	83	33	296
	4	73	62	65	90	2	292
	5	154	127	9	1	1	292
	All	293	293	293	295	296	1471

cross-section's return distribution. Therefore, the same bond can be evaluated by multiple benchmarks: the rating basis portfolio, duration basis portfolio, downside risk basis portfolio, and so on. And different benchmarks are exposed to different sources of risks. A convenient evaluation of bond risk premium sources is the combined weight.

Moreover, the current literature for the cross-section of corporate bond returns[2] is mostly about bond portfolios. Most bond pricing studies are established for rating-sorted and duration-sorted portfolios. Similar to equities, these characteristic-sorted benchmark portfolios possess stabler economic-driven dynamics than individual counterparts. In our approach, the combination weights can be used to identify the underlying economic sources (credit rating, downside risk, return reversal, etc.). In short, our approach consists of two steps: First, we obtain the combination weights by minimizing the pricing error objective. Second, we combine different basis portfolio returns into one single benchmark for pricing tasks.

[2]Related work includes Lin et al. (2011), Bai et al. (2019a), and Gao et al. (2020)

1.3 Empirical Overviews

Our data consist of 589,528 bond-month observations in the empirical study, covering 19,782 unique bonds, from 1976 to 2017. First, we find credit rating, downside risk, and short-term reversal basis portfolios are economically and statistically significant sources in the benchmark combination model (BCM). Second, our benchmark evaluation approach outperforms widely used factor models (five-factor model of Fama and French (1993) and the four-factor model of Bai et al. (2019a)) in pricing individual corporate bond returns. Third, we find that the sorted-portfolio returns' conditional expectation can be a good forecast for the underlying individual bond returns. Our BCM delivers substantial out-of-sample predictability for individual corporate bond returns. Finally, our model-implied long-short strategy, which is not explained by the five-factor model of Fama and French (1993), delivers an 11.88% annualized average return and a 2.57 Sharpe ratio.

The remainder of the paper is organized as follows. Section 1.4 lists the relevant literature and the position of our work. Section 2 demonstrates our BCM model and its connection to empirical asset pricing. We introduce the data in Section 3 and empirical study design in Section 4. In Section 5.1, we show how to decompose the individual corporate bond returns into basis portfolio returns and confirm the source of risk premia. We present empirical evidences for the pricing and prediction performance in Section 5.2 and 5.3. Section 6 summarizes the paper.

1.4 Related Literature

This paper contributes to the literature in several important ways. The foremost contribution is pricing corporate bond returns. Fama and French (1993) propose a five-factor model to price the joint cross-section of equity and corporate bond returns. Bai et al. (2019a) introduce common risk factors based on the prevalent risk characteristics of corporate bonds, including downside risk, credit risk, liquidity risk, and short-term reversal. The proposed factor models in the previous two papers are tested with corporate bond portfolios. Recently, Kelly et al. (2019) try to price individual corporate bond returns with the IPCA method proposed in Kelly et al. (2019). Unlike the traditional time-series regression approach to estimate factor loadings, the IPCA incorporates information on corporate bond characteristics to estimate the factor loadings. Additionally, a large body of literature exists on the cross-section of corporate bond returns. Lin et al. (2011) find liquidity risk is an important determinant, and Chung et al. (2019) document the volatility risk and Huang et al. (2020) confirm the downside risk. Gao et al. (2020) find media coverage is negatively associated with firms' cost of debt.

Second, we contribute to the literature on characteristic-sorted benchmarks. Daniel et al. (1997) provide a benchmark evaluation framework for individual equity returns, which dependently sorts on size, value, and momentum. Portfolios are commonly sorted on a single characteristic, two characteristics, and even three characteristics. However, high-dimensional sorting is difficult as questioned in Cochrane (2011). We provide a solution to the high-dimensional sort difficulty.

Third, our paper is linked to the area of corporate bond return prediction. Lin et al. (2014) document the forward rate, liquidity factor, and credit spread significantly contribute to the predictive power of corporate bond returns, especially through combination forecasts. Chordia et al. (2017) also find predictability facts with equity characteristics, such as profitability and asset growth, and Bai et al. (2019b) confirm the predictability of return volatility and skewness. Lin et al. (2018) find predictability facts for corporate bond portfolio returns with an iterated combination approach. Recent working papers, including Bali et al. (2020) and Bredendiek et al. (2019), predict public individual corporate bond returns or create optimal portfolios using both bond and equity characteristics.

Finally, this paper is related to rising literature on applied machine learning in empirical asset pricing. Gu et al. (2020) forecast individual equity returns with various machine learning algorithms, and Feng et al. (2020) find various portfolio returns are predictable and adopt a forecast combination for individual stock returns. Bianchi et al. (2020) and Feng et al. (2020) applied machine learning and deep learning to investigate the return predictability of U.S. treasury bond returns through revised and real-time macroeconomic variables. A recent working paper by Guo et al. (2020) finds machine learning predictability of corporate bond returns, using yield predictors that capture the yield curve information.

2 Methodology

2.1 Asset Pricing via Optimization

Asset pricing is essentially an optimization problem, which minimizes the pricing error, while constraining the structure of the benchmark returns $\mathcal{X}_{i,t}$ to be economically interpretable:

$$\underset{\mathcal{X}_{i,t}}{\text{Minimize}}\ \mathcal{L} = \sum_{i,t}(r_{i,t} - \mathcal{X}_{i,t})^2, \tag{1}$$

where $r_{i,t}$ denotes the return of asset i at time t, and $\mathcal{X}_{i,t}$ represents the benchmark portfolio.

Financial economists usually assume the benchmark return $\mathcal{X}_{i,t}$ is tradable; in other words, it is a linear combination of all the assets' returns in the cross-section. The combination weights shown in Equation 2 can be super naive, such as equal-weight. However, the researchers usually define the combination weights for time t conditional on asset characteristics up to time $t-1$, as shown in Equation 3:

$$\mathcal{X}_{i,t} = W_i^\intercal r_t, \tag{2}$$

where W_i is the weight vector specified for asset i and r_t is the return vector of all assets in the cross-section:

$$\begin{aligned} W_{t-1} &= [W_{1,t-1}, \cdots, W_{N,t-1}] \\ &= F(Z_{t-1}), \end{aligned} \tag{3}$$

where the sorting function $F(Z_{t-1})$ inputs an $N \times K$ matrix and outputs an $N \times N$ matrix, considering a panel data with N assets and K characteristics. The general formulation of $\mathcal{X}_{i,t}$ doesn't reduce the dimension of the asset returns; on the contrary, it increases the dimension from N the number of assets to $N \times N$ the number of parameters in the weight matrix.

2.2 Factor Model

Financial economists propose CAPM, which decomposes the benchmark $\mathcal{X}_{i,t}$ as the product of two components: the factor loading $\beta_{i,t-1}$ and the market factor f_t, as shown in Equation 4. The market factor is a portfolio of all available assets in the cross-section, managed by the market equity $u_{j,t-1}$, as shown in Equation 5:

$$\mathcal{X}_{i,t} = \beta_{i,t-1} \times f_t \tag{4}$$

$$f_t = \sum_{j=1}^{N} \frac{u_{j,t-1}}{\sum_j u_{j,t-1}} r_{j,t} = \sum_{j=1}^{N} \tilde{u}_{j,t-1} r_{j,t} = \tilde{u}_{t-1}^{\intercal} r_t. \tag{5}$$

Thus, the $W_{i,t-1}$ is formulated as the product of two vectors: the factor loading $\beta_{i,t-1}$ and the factor composition $\tilde{u}_{t-1}$. In practice, the factor composition $\tilde{u}_{t-1}$ is pre-specified. So, the parameter needing estimation is only β_{t-1}, the factor loading of dimension N. For a three-factor model, $3 \times N$ number of parameters need estimation, and $5 \times N$ parameters for a five-factor model. On the one hand, the factor model reduces the number of parameters that need estimation. On the other hand, the tradable factors are interpreted as proxies for some common risk factors in the cross-section of asset returns.

2.3 Univariate-Sorted Benchmark Model

Here, we demonstrate the simplest case, the univariate-quintile sorting on characteristic k, and we have five basis portfolios on characteristcs k. All the assets are sorted on their value of characteristic k from low to high. Those assets positioned below 20% are grouped in bucket one, those between 20% to 40% are grouped in bucket two, and so on. We denote the buckets in $s \in \{1, 2, 3, 4, 5\}$. Those bonds in the same bucket have the same s, and those bonds in different buckets have different s. So, we reduce the dimension from N to S, where $S = 5$.

As shown in Equation 6, the main assumption is that the assets in the same sorting bucket s share the same weight $\widetilde{W}_{s,t-1}^{(k)}$. For a single asset i, we look for its belonging bucket s, regarding characteristic k, and then assign the $\widetilde{W}_{s,t-1}^{(k)}$ to $W_{i,t-1}^{(k)\intercal}$:

$$\begin{aligned}
\mathcal{X}_{i,t} &= W_{i,t-1}^{(k)\intercal} r_t \\
&= \widetilde{W}_{s,t-1}^{(k)\intercal} r_t && (6) \\
&= R_{k,s,t} && (7)
\end{aligned}$$

$$\widetilde{W}^{(k)}_{s,t-1}[j] = \begin{cases} \tilde{u}_{j,t-1} & \text{if } j \in \text{ bucket } s, \\ 0 & \text{if } j \notin \text{ bucket } s, \end{cases} \tag{8}$$

where k is fixed, s denotes the sorting bucket of asset i, $W^{(k)}_{i,t-1}$ is replaced by $\widetilde{W}^{(k)}_{s,t-1}$ for all i in bucket s, and $R_{k,s,t}$ is the benchmark return of bucket s sorted on characteristic k.

In practice, we use the equal (value) weight within the sorting bucket, and the $R_{k,s,t}$ is the equal- (value-) weighted return of the basis portfolio s. In Equation 8, the equal-weighted basis portfolio uses $\tilde{u}_{j,t-1} = \frac{S}{N}$, whereas the value-weighted basis portfolio uses the market capitalization as $u_{j,t-1}$. We have no parameter to estimate, because of the non-parametric nature of sorting.

2.4 Benchmark Combination Model

The proposed BCM model benefits from a large number of characteristics. We combine the K number of basis portfolios into one benchmark $\mathcal{X}_{i,t}$, and the combination weight is $\omega = [\omega_1, \omega_2, \cdots, \omega_K]^\intercal$:

$$\begin{aligned}
\mathcal{X}_{i,t} &= \omega_1 \widetilde{W}^{(1)\intercal}_{s,t-1} r_t + \omega_2 \widetilde{W}^{(2)\intercal}_{s,t-1} r_t + \cdots + \omega_K \widetilde{W}^{(K)\intercal}_{s,t-1} r_t & (9)\\
&= \sum_{k=1}^{K} \left(\omega_k \widetilde{W}^{(k)\intercal}_{s,t-1} r_t \right) & (10)\\
&= \left(\sum_{k=1}^{K} \omega_k \widetilde{W}^{(k)\intercal}_{s,t-1} \right) r_t. & (11)
\end{aligned}$$

The Equation 11 is rewriting Equation 2, but the $W_i^\intercal$ is represented by $\sum_{k=1}^{K} \omega_k \widetilde{W}^{(k)\intercal}_{s,t-1}$. We need to estimate the combination weight ω; thus, the number of parameters that need estimation is K.

Also, we can write the BCM in a concise way:

$$\mathcal{X}_{i,t} = \sum_{k=1}^{K} \omega_k R_{k,s,t}, \tag{12}$$

where $R_{k,s,t}$ is the basis portfolio return of asset i based on characteristic k, and the notation i is embedded in s.

2.5 Estimation via Optimization

The BCM model transforms the optimization problem in Equation 1 to Equation 13 with constraints in Equation 14. We expect the combination weight to be non-negative and sum to unity. Thus, a clear pattern emerges in which the important basis portfolios have large weights and the trivial basis portfolio have zero weights. Also, we consider an unconstrained version for a robustness check. [3]

$$\underset{\{\omega_1,\omega_2,\cdots,\omega_K\}}{\text{Minimize}}\ \mathcal{L} = \sum_{i,t}(r_{i,t} - \sum_{k=1}^{K}\omega_k R_{k,s,t})^2 \tag{13}$$

$$\text{with} \sum_{k=1}^{K}\omega_k = 1, \omega_k \geq 0 \tag{14}$$

The formulation in Equation 13 is a cousin of the factor model and the univariate-sorted benchmark model. The difference between the three models is that they impose different assumptions to make the weight W interpretable and simplified.

2.6 Return Prediction

We extend the pricing model in Equation 12 to a predictive framework:

$$E_t(r_{i,t+1}) = E_t(\mathcal{X}_{i,t+1}) = \sum_{k=1}^{K}\omega_{k|t}E_t(R_{k,s,t+1}), \tag{15}$$

where $E_t(r_{i,t+1})$ is the conditional expectation of the asset i's return at time $t+1$ based on information up to time t, $w_{j|t}$ is the out-of-sample estimation of the combination weights, and $E_t(R_{k,s,t+1})$ is the conditional expectation of basis portfolio returns.

The out-of-sample weights can be estimated with the constrained optimization in Equation 13. As for the $E_t(R_{k,s,t+1})$, it is essentially a prediction for the basis portfolio returns. We formulate it in a general predictive model:

$$E_t(R_{k,s,t+1}) = g_{k,s|t}(\widetilde{Z}_{k,s,t}, x_t), \tag{16}$$

[3] The weight can be estimated in alternative ways. The finance literature has a keen on Fama-Macbeth regression, but adding constraints on the regression coefficients is hard. The unconstrained optimization is reported in this paper for a robustness check. Its empirical performance may be better than the constrained optimization, but the interpretability is undermined. Another way is through panel regression, which is equivalent to the unconstrained optimization.

where two groups of predictors are included: $\widetilde{Z}_{k,s,t}$ is the idiosyncratic signals of the sorted portfolios, and x_t is the systematic signals. [4]

The predictive function $g(\cdot)$ can be any predictive model, including machine learning. We take the $E_t(R_{j,k,t+1})$ from some predictive model, and then refer to the predictive framework in Equation 15. We short-list five predictive models. Two are traditional models: historical average and mean combination, seen in Lin et al. (2014). The others are machine learning models: principal component regression, Lasso, and random forest, seen in Gu et al. (2020). Our goal is not to do a horse-race among many predictive models, but to show our benchmark combination model framework is robust to out-of-sample returns prediction.

3 Empirical Data

3.1 Corporate Bond Returns Sample

The corporate bond observations are collected from four databases: the Lehman Brother Fixed Income (LBFI) database, DataStream, the National Association of Insurance Commissioners (NAIC) database, and the Trade Reporting and Compliance Engine (TRACE) database. We combine the data from these sources to get a large individual corporate bond sample. If we find duplicate observations between multiple sources, we keep only one of them by a priority ranking. The ranking from high priority to low priority is TRACE, NAIC, LBFI, and DataStream. We prefer using the transaction-based return data (TRACE) to using the return data based on quotes and matrix calculations (LBFI). The TRACE data start in 2002 and have a short history, whereas the LBFI extends our sample to early 1973.

With the five sources, we can compute corporate bond returns and various characteristics. The monthly corporate bond return at time t is calculated as follows:

$$R_t = \frac{(P_t + A_t) + C_t - (P_{t-1} + A_{t-1})}{P_{t-1} + A_{t-1}}, \tag{17}$$

where P_t is the price, A_t is the accrued interest, and C_t is the coupon payment at time t. We adjust

[4]Using both of the idiosyncratic signals and the systematic signals for returns prediction is common practice. Such examples can be seen in Gu et al. (2020) and Feng et al. (2020).

the raw returns to excess returns by subtracting the three-month Treasury bill rate. Our sample excludes any bonds with embedded options. We remove any bond observation whose time-to-maturity is less than two years or longer than 30 years. Summary statistics of the corporate bond returns sample are shown in Table A.1 and Table A.2.

3.2 Bond Characteristics, Macro Predictors, and Sorted Portfolios

We have 20 corporate bond characteristics for individual bonds and 20 macro predictors, described in Appendix B.1 and Appendix B.2, respectively. The corporate bond characteristic sample covers four primary categories: the fundamental characteristics (e.g., rating, duration), the return-distributional characteristics (e.g., variance, downside risk), past returns (e.g., short-term reversal, momentum), and covariance with common factors (e.g., beta on TERM, beta on DEF).

The Mergent FISD database provides corporate bond issue information and many characteristics, such as issue size, issue date, maturity date, coupon interest rate, credit rating, and SIC code. Besides the bond characteristics available from Mergent FISD, we calculate more corporate bond characteristics based on past returns. First, following those return-distribution characteristics proposed in Bai et al. (2019a), and Bai et al. (2019b), we have added downside risk proxied by the 5% VaR (value-at-risk), variance, skewness, and excess kurtosis of the returns, which are estimated using the rolling sample of the past 36 months. Second, we have created five past return-related characteristics, including short-term reversal (1-1 month), momentum (2-6 month, and 2-12 month), and long-term reversal (13-24 month, and lagged 13-36 month). Third, following the risk exposure to the corporate bond's risk factors in Fama and French (1993) and Gebhardt et al. (2005), we estimate the multivariate betas of individual bonds regarding a five-factor model, using the rolling sample of the past 36 months. The five-factor model includes the Fama-French three factors (MktRf for the market, SMB for size, HML for value), plus two bond-related factors (term factor and default factor). We also retain the residual variances from the five-factor model regression.

In addition to 20 corporate bond characteristics, we have added another group of predictors, 20 macroeconomic indicators. These macroeconomic indicators cover three main categories: the macroeconomic indicators (e.g., CPI index, three-month Treasury bill rate), the corporate bond market variable (e.g., term spread, corporate bond market return), and the equity market variable (e.g., S&P 500 index returns, S&P 500 index earnings-to-price ratio). The macroeconomic indicators serve as the systematic signal in Equation 16.

Finally, based on each characteristic, we construct the basis portfolios (univariate-sorted quintile portfolios) and rebalance them every month. We calculate the basis portfolio returns as the equal-weighted average of the underlying assets' returns. Also, the characteristics of the basis portfolios are the equal-weighted average of the underlying assets' characteristics. The basis portfolio characteristics serve as the idiosyncratic signal in Equation 16.

4 Empirical Design

4.1 Pricing Measure

Once the weight is estimated, we use Equation 12 to construct our model-implied benchmark for each individual corporate bond return observations. We use the pricing performance measure Total R^2 and Predictive R^2 proposed by Kelly et al. (2019).

$$\text{Total } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t} - \sum_{k=1}^{K}\hat{\omega}_k R_{k,s,t}\right)^2}{\sum_{i,t} r_{i,t}^2}, \tag{18}$$

where $\hat{\omega}_k$'s are the estimation for weights. The Total R^2 represents the fraction of realized return variation explained by the contemporaneous factor realizations, aggregated over all assets and all periods. For in-sample pricing, the $\hat{\omega}_k$'s are estimated with the whole sample from year 1998 to 2017, whereas for out-of-sample pricing, they are estimated with information up to time $t-1$. Thus, we denote if as $\hat{\omega}_k = \hat{\omega}_{k|t-1}$.

$$\text{Predictive } R^2 = 1 - \frac{\sum_{i,t}\left(r_{i,t} - \sum_{k=1}^{K}\hat{\omega}_k \bar{R}_{k,s,t-1}\right)^2,}{\sum_{i,t} r_{i,t}^2} \tag{19}$$

where $\bar{R}_{k,s,t-1}$ denotes the moving average of the basis portfolio returns up to time $t-1$, and $\hat{\lambda}_{t-1}$ is the moving average of the tradable factor risk premium up to time $t-1$. The Predictive R^2 represents the fraction of realized return variation explained by the model-implied expected returns. For the comparison of factor models, we have the same design for in-sample and out-of-sample estimation and follow the calculation in Kelly et al. (2019).

4.2 Prediction Measure

To evaluate the predictability of BCM model, we report the out-of-sample R^2 for predicting basis portfolio returns. We calculate the prediction error for basis portfolio return $\hat{R}_{k,s,t}$ by comparing with its 20-year moving average $\bar{R}_{k,s,t}$.

$$R^2_{OOS,k,s} = 1 - \frac{\sum_t (R_{k,s,t} - \hat{R}_{k,s,t})^2}{\sum_t (R_{k,s,t} - \bar{R}_{k,s,t})^2}, \quad \text{for } k \text{ in } 1, \cdots, K, \text{ and } s \text{ in } 1, \cdots, S. \tag{20}$$

In particular, we have aggregated the prediction errors for the cross-section sorted on each characteristic k.

$$R^2_{OOS,k} = 1 - \frac{\sum_{s=1}^{S=5} \sum_t (R_{k,s,t} - \hat{R}_{k,s,t})^2}{\sum_{s=1}^{S=5} \sum_t (R_{k,s,t} - \bar{R}_{k,s,t})^2}, \quad \text{for } k \text{ in } 1, \cdots, K. \tag{21}$$

We evaluate the predictability of individual bond returns with the below out-of-sample R^2. The baseline prediction $\bar{r}_{i,t}$ is the 20-year moving average return for the corresponding credit rating basis portfolio.[5]

$$R^2_{OOS} = 1 - \frac{\sum_{i,t} (r_{i,t} - \hat{\mathcal{X}}_{i,t})^2}{\sum_{i,t} (r_{i,t} - \bar{r}_{i,t})^2}. \tag{22}$$

4.3 Fama-Macbeth Testing

The above Total, Predictive, and out-of-sample R^2's are aggregate measures over all time periods. However, the drawback for such aggregate measures is they cannot reflect the performance variation over time and can be dominated by extremely volatile periods. We follow Fama and MacBeth (1973) and design the Fama-Macbeth type measures for pricing and prediction. In particular, we utilize the large cross-section and aggregate the information for each time period to obtain the R^2_t. Therefore, we are able to perform the Fama-Macbeth test on the average performance for the time series $\{R^2_t\}_{t=1}^T$ with the below hypothesis.

$$\text{H0: } \overline{R^2} > 0\, ; \text{H1: } \overline{R^2} \leq 0.$$

[5] For example, a AAA bond uses the average returns of AAA portfolio.

For example, we can calculate the periodical version of Total R_t^2 as

$$\text{Total } R_t^2 = 1 - \frac{\sum_i \left(r_{i,t} - \sum_{k=1}^{K} \hat{\omega}_k R_{k,s,t} \right)^2}{\sum_i r_{i,t}^2}. \tag{23}$$

Given the flexibility of designing the Fama-Macbeth test, we can apply the two-sample t-test for the performance comparison. We can test the below hypothesis to compare the average difference for performance measures.

$$\text{H0: } \overline{R_\Delta^2} = 0 \text{ ; H1: } \overline{R_\Delta^2} \neq 0,$$

where $R^2_{\Delta,t} = R^2_{1,t} - R^2_{2,t}$. We can perform this comparison test on two different methods or two different samples.

4.4 Estimation and Prediction

First, we need to estimate the basis portfolio weights for the Benchmark Combination Model. We implement the constrained regression for realized returns in Equation 13. For out-of-sample estimation, we update the model annually using the rolling window of the past 60 months. To obtain the significance of portfolio weights, we implement 1000 bootstrap samples and report the non-negative estimation frequency for each benchmark to measure weight significance. Variables estimated with zero coefficients are useless for the objective in the bootstrap sample. A variable that appeared more than 90% is deemed significant. We present the bootstrap significance levels by dark and light color in Table 2.

Second, we also perform the return prediction for each basis portfolio. In particular, we train the prediction model by the time-series modeling instead of a pooled modeling. The out-of-sample period is January 1998 to December 2017, and the training window is the rolling window of the past 20 years. We update the model annually and predict all 12 observations for the next year. The prediction design follows He et al. (2021). In addition to the moving average, we consider four major forecasting methods, including the mean combination, Principal Component Regression, Lasso, and Random Forest.

5 Empirical Findings

5.1 Returns Decomposition

As discussed in Section 2, the Benchmark Combination Model is an alternative approach to the factor model to provide an expected return and reveal the risk premia sources. The first column in Table 2 reports the in-sample results, using all return observations from 1998 to 2017. We regress individual bond realized returns on their corresponding basis portfolios. We find Credit Rating, Short-term Reversal, 12-month Momentum, and Downside Risk are important sources by the bootstrap inference. In addition to rating, our findings on short-term reversal and downside risk are consistent with Bai et al. (2019a) and Bai et al. (2019b).

Under different economic conditions, the basis portfolio importance can be different in the in-sample analysis. We split the 20 years from 1998 to 2017 into the economic boom and recession with NBER business-cycle indicators.[6] We find Credit Rating is always strong. However, Duration and Short-Term Reversal are more important in the economic boom than in the recession. Downside Risk and 12-Month Momentum are significantly pricing sources in the recession but not in the economic boom.

The remaining columns report the out-of-sample results of combination weights. The combination weights in each year are estimated using the past five-year rolling-window sample. First of all, most weights are zero due to the constraint estimation. Second, Short-Term Reversal is strong in the entire sample. Third, credit Rating becomes dominant since the year 2002. Finally, Downside Risk is only important right after the 2008-2009 financial crisis.

5.2 Pricing Performance

5.2.1 Pricing Performance

Table 3 reports the pricing performance of two BCM models and two common factor models. BCMc is the constrained model, and BCMu is the unconstrained one. The factor models follow the five-factor model (MKT, SMB, HML, TERM, DEF) in Fama and French (1993), and the four-

[6] See https://www.nber.org/research/business-cycle-dating.

factor model (CBMKT, CRF, DRF, STR) in Bai et al. (2019a).[7] Panel A reports in-sample Total R^2, where the BCM combination weights are estimated from 1998 to 2017, and the factor model betas are estimated with the entire history of each bond.[8] Panel B reports out-of-sample Total R^2, where the BCM combination weights are dynamically updated with the past five-year data, and the factor model betas are estimated recursively with the past-three-year samples for each bond.[9]

We can only find similar in-sample Total R^2 between the BCM models and factor models because of the winsorized beta estimates for the factor model. The BCM models are more robust than the factor model for pricing individual assets with limited observations. The BCM models significantly outperform factor models regarding in-sample Fama-Macbeth $\overline{R^2}$, which shows the BCM performance is more stable over time. For out-of-sample Total R^2, as shown in Panel B of Table 3, the BCM dramatically outperforms the factor models that even underperform the zero return baseline. Besides, we present the Predictive R^2 reported in Panel C. With pre-estimated combination weights and basis portfolio return forecasts, the Predictive R^2 describes the individual bond risk premia with a pure forecasting perspective. Negative numbers in Predictive R^2 imply a model underperforms the zero average return baseline. We find both BCM models generate significantly positive Predictive R^2, while factor models generate insignificant or even negative numbers. Finally, these results are robust for investment-grade and non-investment-grade bonds and suggest the BCM model as a better alternative to measure the risk premium for individual bonds.

5.2.2 Model Comparison

We want to show our BCM models significantly outperform factor models and average returns in describing the cross-section for different cases. Given the large cross-section of the panel data, the Fama-Macbeth test can be flexible and useful for multiple comparisons. We consider a two-sample Fama-Macbeth test to compare the $\overline{R^2}$ between the BCM models and factor models. In Table 4, we report the average performance difference, $\overline{R^2_\Delta}$, of the model in the row over the column model. Both BCM models deliver significantly larger $\overline{R^2_\Delta}$ than FF5 and BBW4 in all pricing metrics and

[7] We follow Bai et al. (2019a) and construct our Corporate Bond Market, Downside Risk, Credit Rating, and Short-Term Reversal using on our sample.

[8] Given high volatility on individual bond betas, we have winsorized the beta estimates by 5% and 95% quantiles in the cross-section.

[9] We estimate the out-of-sample individual bond betas with the past 36-month returns sample. For missing returns, we use the corresponding 5×5 rating-duration portfolio return for the imputation.

different subsamples.

Besides, we present the pricing performance of the BCMc model on each bivariate-sorted portfolio on Credit Rating and Duration. We report the Predictive R^2 in Table 5. In Panel A, we find the BCMc model produces significantly positive numbers for long-duration (and high-rating) bonds. We also report the positive performance of moving-average returns for the bivariate-sorted portfolio over the zero average return baseline. Finally, we report our BCMc model's significant outperformance over moving-average returns for bivariate-sorted portfolio in Panel C.

5.2.3 Private versus Public Bonds

There are recent studies (Chordia et al., 2017, Choi and Kim, 2018, and Bali et al., 2020) about the equity characteristics information in pricing the cross-section of corporate bond returns. Unfortunately, such attempts are only feasible for those corporate bonds issued by publicly listed firms. For bonds issued by private firms, access to the equity characteristics is highly limited. All the above studies only consider corporate bonds issued by publicly listed firms in their sample. The usefulness of equity characteristics can be due to the sample selection for publicly listed firms. In our extensive bond sample, however, more than 75% of bond-return observations are from private firms.

We study this relationship with a different angle: whether public and private bonds have the same pricing difficulty (signal-to-noise ratio) when using the same predictors without equity characteristics. In Table 6, Panels A and B report the Total R^2 for private bonds, public bonds, and their difference, whereas Table 6 Panel C reports the Predictive R^2. We find the private bonds have significantly smaller Total R^2 and Predictive R^2 than public bonds for the Fama-Macbeth testing. Without using any equity information, we find the public bonds have smaller pricing errors than private bonds. These conclusions are all true using our BCM models, where factor models only show negative differences for the out-of-sample analysis. Therefore, there is a significant sample selection bias between public and private bonds.

5.3 Prediction Performance

5.3.1 Portfolio Return Predictability

First of all, we want to present the return predictability evidence for various basis portfolios. We report the out-of-sample R^2 in Equation 20 in Table 7 for methods listed in Section 4.4. A category noted "All" reports the aggregate out-of-sample R^2 (Equation 21) of basis portfolios for each characteristic. Table 7 reports the predictability evidence for those most important characteristics in our findings, such as Credit Rating, Duration, Age, Short-Term Reversal, and Downside Risk. We have reproduced consistent results in Lin et al. (2014) and Lin et al. (2018), such that the Mean Combination can predict the Credit Rating or Duration basis portfolio returns. Furthermore, we find machine learning methods can substantially increase the forecast performance, which implies the potential nonlinear predictability for the predictor-return structure.

Also, we have extended the return predictability studies to all 20 groups of basis portfolio returns. We have provided Table A.7 in the appendix, which reports the aggregate predictability for all basis portfolio returns regarding each characteristic. For all basis portfolios, we find highly promising predictability evidence over the average benchmark. In particular, the traditional Mean Combination delivers robust predictability for all basis portfolios. Moreover, machine learning methods, including Lasso, PCA regression, and Random Forest, deliver extremely robust predictability evidence. These are all new findings to study the return predictability for bond basis portfolios.

5.3.2 Predicting Individual Bond Returns by Basis Portfolio Prediction

We start to present using the predictability of basis portfolios for individual bond returns. Table 8 reports the out-of-sample R^2_{OOS} of predicting individual bond returns based on Equation 22. It reports the results for different subsamples, including investment-grade bonds, non-investment-grade bonds, and the overall sample.

For the average return predictions, the Credit Rating basis portfolio reports zeros for columns of IG, NIG, and Overall, because, in Equation 22, both the prediction $\hat{r}_{i,t}$ and the baseline model $\bar{r}_{i,t}$ are average returns of the Credit Rating basis portfolio. The Mean Combination, Lasso, PCA regression predictions, and Random Forest offer a tremendous improvement over the average baseline. In Table 7, we find the machine learning forecasts for basis portfolio returns are way more powerful than the average returns. Table 8 further shows these predictability power on basis portfolio returns can be used to predict individual bond returns.

5.3.3 Predicting Individual Bond Returns by Benchmark Combination Model

We now show how to create a return forecast based on the benchmark combination model's pricing kernel. Step one, we can create the return forecast for the BCM model components, various basis portfolios. We combine these return forecasts with the BCM combination weights to obtain a prediction model in step two. With the dynamic weights in Table 2, we construct a combination prediction (or ensemble prediction) for individual bond returns.

Panel A of Table 9 reports the benchmark combination forecast's out-of-sample performance. Taking Lasso as an example, we have Lasso return predictions for various basis portfolios and combine these Lasso return forecasts with the BCM weights. Finally, we report the out-of-sample R^2_{OOS} of predicting individual bond returns by the weighted combination Lasso forecast, labeled "c-Lasso" in the table. The traditional Mean Combination delivers slightly positive results, while Lasso, PCA regression, and Random Forest provide substantial out-of-sample predictability.

Panel B reports the BCM forecast with unconstrained weights for a robustness check (labeled "u-Model" in the table). In Panel C, the naive mean benchmark combination (equally weighted combination) provides positive results (labeled "m-Model" in the table.). Still, the magnitude is smaller than the optimally-weighted BCM results. We also report results for the factor model, though most numbers are mostly negative. Table A.8 in Appendix reports the subperiod prediction results for individual bond returns. We find the predictability is consistently positive for all subperiods but stronger during and after the 2008-2009 financial crisis.

5.3.4 Investment Performance

Based on the BCM forecasts for individual corporate bond returns, we construct forecast-implied quintile and long-short portfolios. Table 10 reports the performance measures for these long-short portfolios. We have presented the average returns, α's based on the five-factor model

(MKT, SMB, HML, TERM, DEF), t-stat's for the Alpha, and the annualized Sharpe ratio. We find that the averages returns have a substantial increase over the corporate bond market portfolio (investment-grade, non-investment-grade, and overall bonds). These positive investment performances are robust for all methods, including the BCM forecasts using basis portfolio average returns. The α's are positive and very close to the average returns, indicating the five-factor model cannot explain the long-short strategy returns. The out-of-sample portfolio cumulative log returns are shown in Figure A.1, where we find pretty clear cross-sectional differences.

6 Conclusion

This paper shares an alternative approach to analyzing the individual corporate bond risk premia. The proposed benchmark combination model (BCM) is related to the linear factor asset pricing model and the basis portfolio benchmark evaluation of Daniel et al. (1997). We attempt to solve the high-dimensional sort difficulty questioned in Cochrane (2011), with a linear combination of basis portfolios. We find three main drivers of individual corporate bond risk premia: Credit Rating, Short-Term Reversal, and Downside Risk. Empirically, the BCM model outperforms commonly used factor models in pricing individual corporate bonds. Moreover, the BCM model produces substantial evidence of return predictability for individual corporate bonds and bond basis portfolios.

References

Bai, J., T. G. Bali, and Q. Wen (2019a). Common risk factors in the cross-section of corporate bond returns. *Journal of Financial Economics 131*(3), 619–642.

Bai, J., T. G. Bali, and Q. Wen (2019b). Do the distributional characteristics of corporate bonds predict their future returns? Technical report, Georgetown University.

Bali, T. G., D. Huang, F. Jiang, and Q. Wen (2020). The cross-sectional pricing of corporate bond using big data and machine learning. Technical report, Georgetown University.

Bianchi, D., M. Büchner, and A. Tamoni (2020). Bond risk premiums with machine learning. *The Review of Financial Studies*.

Bredendiek, M., G. Ottonello, and R. I. Valkanov (2019). Corporate bond portfolios and asset-specific information. Technical report, Vienna Graduate School of Finance.

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

Chordia, T., A. Goyal, Y. Nozawa, A. Subrahmanyam, and Q. Tong (2017). Are capital market anomalies common to equity and corporate bond markets? an empirical investigation. *Journal of Financial and Quantitative Analysis 52*(4), 1301–1342.

Chung, K. H., J. Wang, and C. Wu (2019). Volatility and the cross-section of corporate bond returns. *Journal of Financial Economics 133*(2), 397–417.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance 66*(4), 1047–1108.

Cochrane, J. H. and M. Piazzesi (2005). Bond risk premia. *American Economic Review 95*(1), 138–160.

Daniel, K., M. Grinblatt, S. Titman, and R. Wermers (1997). Measuring mutual fund performance with characteristic-based benchmarks. *The Journal of finance 52*(3), 1035–1058.

Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics 33*(1), 3–56.

Fama, E. F. and J. D. MacBeth (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy 81*(3), 607–636.

Feng, G., A. Fulop, and J. Li (2020). Real-time macro information and bond return predictability: Does deep learning help? Technical report, City University of Hong Kong.

Feng, G., J. He, X. He, and N. G. Polson (2020). Deep learning for predicting asset returns. Technical report, City University of Hong Kong.

Feng, G., N. Polson, and J. Xu (2020). Deep learning in characteristics-sorted factor models. Technical report, City University of Hong Kong.

Gao, H., J. Wang, Y. Wang, C. Wu, and X. Dong (2020). Media coverage and the cost of debt. *Journal of Financial and Quantitative Analysis 55*(2), 429–471.

Gebhardt, W. R., S. Hvidkjaer, and B. Swaminathan (2005). The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics 75*(1), 85–114.

Gu, S., B. Kelly, and D. Xiu (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies 33*(5), 2223–2273.

Guo, X., H. Lin, C. Wu, and G. Zhou (2020). Extracting information from the corporate yield curve: A machine learning approach. Technical report, Hunan University.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007). The us treasury yield curve: 1961 to the present. *Journal of monetary Economics 54*(8), 2291–2304.

He, X., G. Feng, J. Wang, and C. Wu (2021). Predicting individual corporate bond returns. Technical report, City University of Hong Kong.

Huang, T., L. Jiang, and J. Li (2020). Downside variance premium and the cross-section of corporate bond returns. Technical report, Fudan University.

Jostova, G., S. Nikolova, A. Philipov, and C. W. Stahel (2013). Momentum in corporate bond returns. *The Review of Financial Studies 26*(7), 1649–1693.

Kelly, B. T., D. Palhares, and S. Pruitt (2019). Modeling corporate bond returns. Technical report, Yale University.

Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*(3), 501–524.

Lin, H., J. Wang, and C. Wu (2011). Liquidity risk and expected corporate bond returns. *Journal of Financial Economics 99*(3), 628–650.

Lin, H., J. Wang, and C. Wu (2014). Predictions of corporate bond excess returns. *Journal of Financial Markets 21*, 123–152.

Lin, H., C. Wu, and G. Zhou (2018). Forecasting corporate bond returns with a large set of predictors: An iterated combination approach. *Management Science 64*(9), 4218–4238.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica: Journal of the Econometric Society*, 867–887.

Pástor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Welch, I. and A. Goyal (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies 21*(4), 1455–1508.

Table 2: Weight (%) for Benchmark Combination

This table reports the weight (in percentage %) for portfolio return combinations. The in-sample results are estimated with 20-year observations from 1998 to 2017. Also, we split the 20 years into the economic boom and recession, with the business cycle indicators downloaded from NBER Business Cycle Dating. The out-of-sample weights are updated annually with past-five-year data. The colors of the table cells are conditional on the positive frequency of weights in bootstrap samples. We refit the weights for each of the 1,000 bootstrap samples. Then, we count the positive frequency of each benchmark among the 1,000 repeats. Finally, we categorize the cells into four sets based on the positive frequency, which is divided by three break points—10%, 50%, and 90%—and color the four sets of cells in blank, light, medium, and dark backgrounds.

	In-Sample			Out-of-Sample																			
	1998-2017	Boom	Recession	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Rating	53	38	59	0	0	0	0	15	34	43	46	44	44	41	69	58	58	58	57	40	47	37	38
Duration	0	17	0	18	24	23	14	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
Maturity	0	3	0	6	0	0	5	5	0	7	7	9	9	21	12	0	0	0	0	0	10	21	39
Age	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Size	0	0	0	20	22	28	33	38	25	5	0	0	0	0	0	0	0	0	0	0	0	0	0
S-term Rev	31	35	5	28	33	29	23	19	38	44	46	47	47	38	15	11	12	15	13	24	41	41	18
Mom 6M	0	0	0	15	9	5	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Mom 12M	5	0	13	3	1	1	0	0	3	1	0	0	0	0	3	7	5	5	8	0	0	0	1
L-term Rev 2Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L-term Rev 3Y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variance	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Downside Risk	11	0	15	0	0	0	0	0	0	0	0	0	0	0	0	21	23	21	21	36	1	0	0
Skewness	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kurtosis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_mkt	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_smb	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_hml	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_term	0	4	0	10	12	15	20	13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_def	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	3	2	1	1	0	0	0	0
Residual Var	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bootstrap Positive Frequency						[0, 100)					[100, 500)					[500, 900)					[900, 1000]		

Table 3: Pricing Performance for Individual Corporate Bonds

This table reports the pricing performance for individual corporate bond returns. The BCMc denotes the benchmark combination model with constraints, and the BCMu denotes the benchmark combination model without constraints. The FF5 denotes the five-factor model in Fama and French (1993) and the five factors are MKT, SMB, HML, TERM, and DEF. The BBW4 denotes the factor model proposed in Bai et al. (2019a). Panel A uses in-sample combination weights for BCMc and BCMu, and in-sample beta estimation for FF5 and BBW4. Panel B uses out-of-sample combination weights for BCMc and BCMu, and out-of-sample beta estimation for FF5 and BBW4. Panel C reports the Predictive R^2. Each of the R^2 columns reports one value of pricing performance, pooling all bonds and time periods. FM-$\overline{R^2}$ reports the time-series mean of $\{R_t^2\}_{t=1}^T$. Also, we have the Fama-Macbeth t-test for the time-series mean, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	IG		NIG		Overall	
	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$
Panel A: Total R^2 %, In-Sample						
BCMc	17.96	16.35***	13.20	11.77***	16.00	15.03***
BCMu	19.25	18.05***	14.07	12.59***	17.12	16.52***
FF5	10.02	-6.30**	17.88	-8.52	13.26	-5.05
BBW4	18.49	8.26***	22.24	7.25***	20.03	9.01***
Panel B: Total R^2 %, Out-of-Sample						
BCMc	17.91	16.77***	12.24	11.34***	15.57	15.17***
BCMu	17.65	17.47***	12.87	11.99***	15.68	15.98***
FF5	-12.60	-10.66***	-14.01	-28.35***	-13.18	-14.26***
BBW4	-6.00	0.45	-8.51	-8.22***	-7.04	-0.80
Panel C: Predictive R^2 %						
BCMc	0.49	1.26***	0.41	1.22***	0.45	1.24***
BCMu	0.72	1.69***	0.51	1.56***	0.63	1.63***
FF5	-0.22	0.35	0.16	0.26	-0.06	0.39
BBW4	-0.43	0.12	0.19	0.21	-0.17	0.22

Table 4: Model Comparison

This table reports the two-sample t-test for the time-series R_t^2 of the model in the rows over the model in the columns. The details of the Fama-Macbeth approach are described in Section 4.3. We report the Fama-Macbeth t-statistics, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	IG		NIG		Overall	
	FF5	BBW4	FF5	BBW4	FF5	BBW4
Panel A: Total R^2 %, In-Sample						
BCMc	22.65***	8.09***	20.29***	4.52***	20.08***	6.02***
BCMu	24.36***	9.79***	21.11***	5.35***	21.57***	7.52***
Panel B: Total R^2 %, Out-of-Sample						
BCMc	27.43***	16.32***	39.69***	19.56***	29.43***	15.97***
BCMu	28.13***	17.02***	40.34***	20.22***	30.24***	16.78***
Panel C: Predictive R^2 %						
BCMc	0.91***	1.14***	0.97***	1.01***	0.85***	1.02***
BCMu	1.35***	1.57***	1.31***	1.35***	1.23***	1.41***

Table 5: Fama-Macbeth Predictive $\overline{R^2}$ % by Rating and Duration Style Box

This table reports the Fama-Macbeth Predictive $\overline{R^2}$ of the 25 cells in the 5-by-5 rating and duration style box. Specifically, we report one number and the t-test for each of the 25 cells. Panel A is for the benchmark combination model's constrained version. Panel B is for the bivariate-sorted benchmark on rating and duration. Panel C is for the two-sample t-test between the two models in Panels A and B. ***, **, and * indicate t-stat significance at the 1%, 5%, and 10% level, respectively.

Rating	AAA	AA	A	BBB	NIG
Panel A: Benchmark Combination Model (Constrained)					
Duration					
1	-0.54	-1.55	0.08	0.31	0.94***
2	0.96	1.66***	1.79***	1.34***	1.46***
3	1.35***	1.31***	1.47***	1.37***	1.18***
4	1.92***	1.87***	1.77***	1.38***	1.47***
5	2.13***	1.49***	1.30***	1.69***	1.12***
Panel B: Bivariate-Sort Benchmark on Rating and Duration					
Duration					
1	0.22	-0.11	-0.74	-0.84	0.35
2	0.76	0.85***	1.03***	0.54*	0.88**
3	0.54**	0.48**	0.75***	0.68***	0.86***
4	0.54*	0.85***	1.00***	0.67***	1.17***
5	1.25**	0.80*	0.18	1.36***	0.44**
Panel C: Difference A Minus B					
Duration					
1	-0.76	-1.44	0.82***	1.15***	0.59***
2	0.19	0.81***	0.76***	0.79***	0.59***
3	0.81***	0.83***	0.72***	0.69***	0.32***
4	1.38***	1.03***	0.77***	0.71***	0.30***
5	0.88***	0.69***	1.12***	0.33*	0.68**

Table 6: Pricing Performance for Private and Public Individual Bonds

This table reports the pricing performance for pricing private bonds and public bonds. Panel A uses in-sample combination weights for BCMc and BCMu, and in-sample beta estimation for FF5 and BBW4. Panel B uses out-of-sample combination weights for BCMc and BCMu, and out-of-sample beta estimation for FF5 and BBW4. Panel C reports the Predictive R^2. We conduct a two-sample t-test with the null hypothesis that the time-series mean of private bond R^2 is equal to that of the public bonds, with the results in the last column. We report the Fama-Macbeth t-statistics, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Private		Public		Difference
	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$	FM-$\overline{R^2}$
Panel A: Total R^2 %, In-Sample					
BCMc	15.17	14.22***	18.84	18.10***	-3.88***
BCMu	16.26	15.60***	20.07	19.93***	-4.33***
FF5	11.29	-6.51**	20.03	-1.11	-5.41**
BBW4	18.69	7.98***	24.67	11.80***	-3.82***
Panel B: Total R^2 %, Out-of-Sample					
BCMc	14.85	14.42***	18.09	18.09***	-3.67***
BCMu	14.97	15.09***	18.15	19.24***	-4.15***
FF5	-15.07	-15.56***	-6.66	-13.44***	-2.12
BBW4	-8.17	-2.04	-3.13	0.42	-2.46
Panel C: Predictive R^2 %					
BCMc	0.45	1.17***	0.49	1.56***	-0.39***
BCMu	0.63	1.56***	0.64	1.98***	-0.42**
FF5	-0.11	0.26	0.10	0.67	-0.41
BBW4	-0.18	0.11	-0.16	0.41	-0.30

Table 7: Predicting Corporate Bond Sorted-Portfolio Returns, Out-of-Sample R^2 %

This table reports the out-of-sample R^2_{OOS} values (%) of corporate bond sorted-portfolio excess returns. For each corporate bond characteristic, five sorted portfolios are updated monthly. The prediction methods are shown in each column, and the out-of-sample prediction baseline is the average of each portfolio's excess returns. Here, we show the results for five groups of portfolios sorted on rating, duration, short-term reversal, and downside risk, respectively. Also, we show aggregated results for the five portfolios in each group, noted as "All." The statistics for buckets 1 to 5 are calculated with Equation 20, and the statistics for "All" are from Equation 21. The results for other benchmark predictions are reported in Appendix Table A.7.

Characteristic	Bucket	Mean	Lasso	PCA	Random Forest
Rating	1	5.33	20.36	23.65	31.26
	2	2.53	7.99	6.53	7.65
	3	4.32	7.58	14.42	10.62
	4	4.33	14.59	16.38	7.30
	5	0.34	5.63	11.80	-5.39
	All	2.17	8.88	13.52	3.19
Duration	1	3.99	18.50	20.84	9.55
	2	4.73	18.98	21.64	13.29
	3	3.75	16.07	17.83	11.43
	4	3.07	17.09	14.35	11.94
	5	2.82	14.53	17.27	6.07
	All	3.35	16.24	17.43	9.54
Age	1	4.81	15.19	18.17	12.40
	2	3.01	12.96	16.76	11.45
	3	3.51	12.63	17.53	13.71
	4	3.04	19.92	16.25	8.33
	5	4.05	24.77	21.19	10.87
	All	3.60	17.76	18.03	11.09
Short-term Reversal	1	3.07	22.24	18.72	15.01
	2	6.10	17.16	21.62	20.61
	3	2.87	26.04	12.90	19.37
	4	3.59	15.37	14.56	10.96
	5	3.04	9.49	13.29	11.45
	All	3.50	17.67	16.32	14.75
Downside	1	2.37	9.42	16.44	8.19
	2	3.45	17.15	16.42	9.63
	3	3.67	15.55	15.27	12.22
	4	4.37	12.00	17.98	16.37
	5	5.01	14.51	21.51	22.65
	All	3.14	12.60	16.68	10.71

Table 8: Predicting Individual Bond Returns with Single Benchmark Forecast, Out-of-Sample R^2_{OOS} %

This table reports the out-of-sample R^2_{OOS} % of predicting individual corporate bond returns with the benchmark portfolio forecasts. For each individual bond, we identify the quintile benchmark it belongs to and take the forecasts for the benchmark portfolio returns as the forecasts for the individual bond returns. In the first row, we identify which rating benchmark each bond belongs to and use the rating portfolio returns forecast to predict individual bond returns. In each row, we use a group of sorting on the single bond characteristic. In the columns, we have five methods to predict portfolio returns: (1) average of the portfolio returns, (2) mean combination forecast, (3) Lasso, (4) PCA regression, and (5) random forest. We report the Fama-Macbeth t-statistics, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	Average		MEAN Combination		Lasso		PCA Regression		Random Forest	
Characteristic	R^2	FM-R^2	R^2	FM-R^2	R^2	FM-R^2	R^2	FM-R^2	R^2	FM-R^2
Rating	0.00	0.00	0.36	0.02	1.41	0.31	1.83	0.58	0.79	0.95*
Duration	0.01	0.13*	0.36	0.13*	1.77	0.54	1.67	0.34	1.06	1.18***
Maturity	0.02	0.12**	0.37	0.15**	1.92	0.87	1.76	0.34	1.06	1.13**
Age	-0.02	-0.03	0.32	-0.02	1.82	0.69	1.61	0.14	1.20	1.20***
Size	-0.01	0.05	0.37	0.10	1.78	0.64	1.55	0.09	0.98	1.16***
Short-term Rev	0.60	1.09***	0.98	1.06***	2.65	1.96***	2.19	1.04	2.36	2.81***
Momentum 6M	-0.02	0.00	0.28	-0.04	1.68	0.85	1.42	0.23	0.78	1.31***
Momentum 12M	-0.03	-0.06	0.28	-0.07	1.96	0.97*	1.21	0.15	0.88	1.35***
Long-term Rev 2Y	0.00	-0.21	0.40	-0.20	1.80	0.47	2.23	0.59	1.53	1.18***
Long-term Rev 3Y	-0.02	-0.14	0.41	-0.13	1.99	0.73	2.08	0.65	1.47	1.13**
Variance	0.06	0.21***	0.46	0.21***	1.42	-0.52	2.22	-0.12	1.81	1.35***
Downside Risk	0.14	0.26***	0.50	0.26***	1.72	0.50	1.97	-0.34	1.54	1.38***
Skewness	0.10	-0.14	0.49	-0.14	2.05	0.47	1.66	-0.63	1.51	1.40***
Kurtosis	-0.01	-0.09	0.37	-0.08	1.73	0.22	1.72	0.24	1.40	1.29***
Beta_mkt	0.00	0.03	0.34	0.02	1.78	0.64	1.71	0.06	1.08	1.04**
Beta_smb	0.00	0.02	0.36	0.03	1.38	-0.40	1.57	-0.13	1.16	1.30***
Beta_hml	0.01	0.03	0.35	0.03	1.36	0.30	1.58	0.30	0.87	0.78*
Beta_def	-0.01	-0.03	0.38	0.00	1.95	0.48	1.30	-0.49	1.21	0.94**
Beta_term	0.01	0.04	0.35	0.02	1.75	0.45	1.43	-0.41	1.33	0.94**
Residual Variance	0.05	0.15***	0.43	0.15**	1.99	0.83	2.10	-0.11	1.78	1.46***

Table 9: Predicting Individual Bond Returns with BCM Forecast, Out-of-Sample R^2_{OOS} %

This table reports the out-of-sample R^2_{OOS} values (%) of predicting individual corporate bond excess returns via the BCM forecast. In Panel A, the prediction to an individual corporate bond excess return is the BCM forecast. The weights are reported in Table 2. In Panel B, the combination weights are unconstrained, as reported in Table A.4. In Panel C, the predictions are similar to those of Panel A but use equal weight in combination weights. Panel D reports the prediction performance of factor models. The out-of-sample R^2_{OOS} baseline prediction is the belonging credit rating portfolio average excess return. We report the Fama-Macbeth t-statistics, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	IG		NIG		ALL	
	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$
Panel A: Benchmark Combination Model, Constrained						
c-Avg.	0.39	0.81***	0.09	0.41***	0.27	0.64***
c-Mean Comb.	0.95	1.25***	0.15	0.69***	0.62	1.02***
c-Lasso	3.08	2.22***	1.50	0.47	2.43	1.75***
c-PCA	2.56	1.19	1.66	0.36	2.19	1.12
c-RF	2.66	2.83***	0.27	0.44	1.68	2.18***
Panel B: Benchmark Combination Model, Unconstrained						
u-Avg.	0.65	1.30***	0.19	0.77***	0.46	1.06***
u-Mean Comb.	1.27	1.85***	0.17	1.11***	0.82	1.54***
u-Lasso	3.22	2.58***	1.27	0.17	2.42	1.94***
u-PCA	2.49	1.06	1.41	-0.06	2.04	0.97
u-RF	2.77	3.05***	-0.12	0.10	1.58	2.27***
Panel C: Benchmark Combination Model, Mean Combination						
m-Avg.	0.15	0.34***	0.01	-0.03	0.09	0.20***
m-Mean Comb.	0.66	0.72***	0.16	0.19	0.46	0.54***
m-Lasso	2.65	1.69***	1.85	0.73	2.32	1.45***
m-PCA	2.44	0.77	1.65	0.42	2.11	0.82
m-RF	2.16	2.12***	0.89	0.81*	1.64	1.80***
Panel D: Factor Model						
FF5	-0.46	-0.34*	-0.08	-0.46*	-0.31	-0.35*
BBW4	-0.68	-0.54***	-0.05	-0.50**	-0.42	-0.50***

Table 10: Prediction Implied Long-Short Strategy Returns

This table reports the summary statistics for out-of-sample long-short portfolio returns. The statistics include expected returns (%), alphas (%) on a five-factor model (MKT, SMB, HML, TERM, DEF), t-stat for the alpha, and annualized Sharpe ratios. ***, **, * indicate the significance of alpha at the 1%, 5%, and 10% level, respectively.

	IG			NIG			ALL		
Method	Mean	α	SR	Mean	α	SR	Mean	α	SR
cbmkt	0.23	0.07	0.66	0.49	0.26*	0.61	0.27	0.10*	0.73
Panel A: Weighted Benchmark Combination									
c-avg.	0.99	0.99***	3.22	1.19	1.07***	1.41	0.94	0.93***	2.57
c-mean comb.	1.05	1.04***	3.29	1.07	0.91***	1.34	0.91	0.89***	2.20
c-Lasso	0.93	0.92***	2.73	1.14	1.02***	1.35	0.82	0.81***	1.81
c-pca	0.76	0.74***	2.25	0.83	0.72***	1.01	0.73	0.71***	1.51
c-rf	0.92	0.87***	2.84	1.12	0.99***	1.55	0.78	0.73***	1.73
Panel B: Weighted Benchmark Combination (Unconstrained)									
u-avg.	0.85	0.75***	2.24	1.30	1.15***	1.36	0.93	0.82***	1.93
u-mean comb.	0.89	0.74***	2.32	1.08	0.84***	1.19	0.86	0.69***	1.74
u-Lasso	0.75	0.67***	2.43	1.17	1.09***	1.64	0.73	0.66***	1.70
u-pca	0.59	0.48***	1.63	0.46	0.30*	0.57	0.62	0.51***	1.25
u-rf	0.71	0.58***	2.05	0.75	0.56***	0.82	0.70	0.55***	1.44
Panel C: Mean Benchmark Combination									
m-avg.	0.98	0.98***	3.43	1.06	1.07***	1.27	0.97	0.96***	3.06
m-mean comb.	1.00	0.98***	3.69	1.03	0.94***	1.27	0.87	0.82***	2.39
m-Lasso	0.85	0.84***	3.13	1.02	0.92***	1.21	0.76	0.75***	2.02
m-pca	0.73	0.71***	2.46	0.91	0.75***	1.07	0.68	0.66***	1.58
m-rf	0.92	0.90***	3.39	1.01	0.85***	1.27	0.81	0.77***	2.19

Appendices

A Additional Figures and Tables

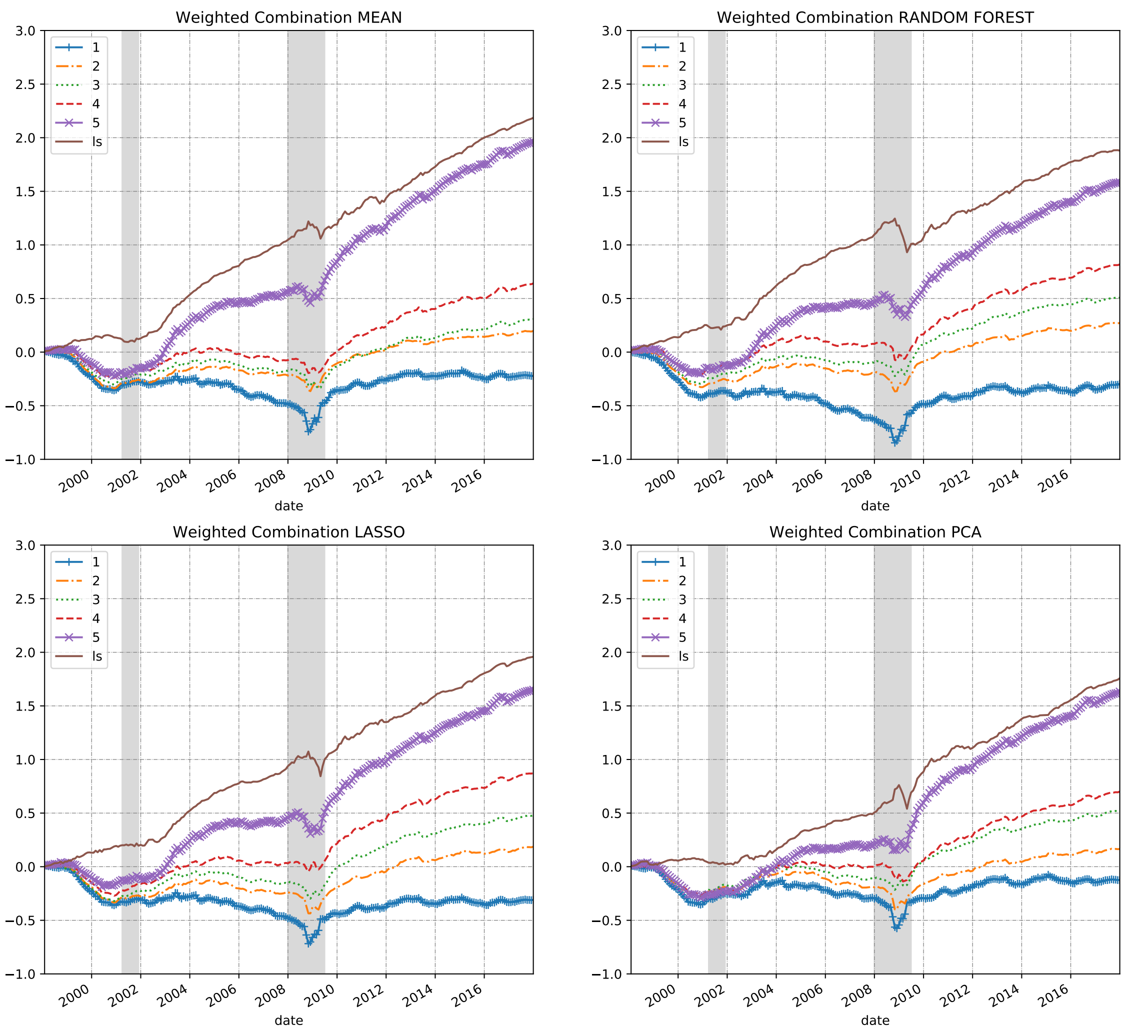

Figure A.1: Cumulative Log Return of Benchmark Combination Forecast & Machine Learning

This figure shows the cumulative log return of portfolios sorted on the out-of-sample forecast of individual corporate bond returns. The assets include both investment-grade and non-investment-grade bonds. We report the results for the weighted BCM forecast of four predictive models (mean, Lasso, PCA, random forest).

Table A.1: Descriptive Statistics

Our final data sample includes 589,528 monthly return observations of 19,782 unique corporate bonds from January 1976 to December 2017. The raw data start in 1973. However, we require a three-year window to initialize risk characteristics such as β_{term}, β_{smb}, and so on. We report TRACE and NAIC together because they are both transaction-based data, and a large proportion of NAIC observations are covered by TRACE.

	All Databases	Lehman	DataStream	TRACE&NAIC
Bond-month observations	589,528	126,710	63,107	399,711
Period coverage	1976-2017	1976-1998	1990-2008	1994-2017
Return - mean (%/month)	0.51	0.73	0.50	0.44
Return - median (%/month)	0.52	0.79	0.61	0.52
Excess Return - mean (%/month)	0.14	0.15	0.18	0.13
Excess Return - median (%/month)	0.22	0.27	0.25	0.18
Rating - mean	5.62	5.18	7.21	5.51
Rating - median	5.00	5.00	7.00	5.00
% of rated that are IG	89.64	93.53	84.20	89.27
% of rated that are NIG	10.36	6.47	15.80	10.73
% of Private bond	78.81	82.66	94.37	75.13
Duration - mean (years)	6.25	5.56	8.67	6.09
Duration - median (years)	5.48	5.32	9.00	5.06
Age - mean (years)	8.32	17.61	6.42	5.68
Age - median (year)	5.37	19.90	5.89	3.97
Amt outst. - mean ($ millions)	10,174.41	708.36	1,861.02	11,441.48
Amt outst. - mean ($ millions)	1,500.00	200.00	1,000.00	2,000.00

Table A.2: Individual Corporate Bond Returns Sample Distribution

Our data sample includes 589,528 monthly return observations of 19,782 unique corporate bonds from January 1976 to December 2017. Panel A reports the percentage distribution of the corporate bond data by credit ratings and maturity length. Panel B reports the percentage distribution of the data by rating and data source.

	AAA	AA	A	BBB	Junk	All
Panel A: By Rating & Maturity						
2	1.87	2.66	5.45	2.78	1.11	13.87
3	1.34	2.17	4.62	2.44	1.00	11.56
4	1.34	2.06	4.61	2.39	0.98	11.37
5	0.79	1.27	3.04	1.78	0.87	7.75
6	0.80	1.16	2.92	1.77	0.86	7.51
7	0.60	1.04	2.54	1.54	0.66	6.39
8	0.59	1.03	2.47	1.56	0.59	6.24
9	0.57	1.01	2.57	1.75	0.62	6.51
10	0.16	0.53	1.07	0.88	0.33	2.96
>10	1.97	3.34	8.93	8.37	3.21	25.83
All	10.02	16.27	38.22	25.27	10.22	100.00
Panel B: By Rating & Data Source						
Datastream	0.13	0.82	3.33	4.57	1.66	10.50
LBFI	1.44	5.99	9.11	5.01	1.43	22.99
NAIC	7.35	3.32	11.84	4.65	1.39	28.55
TRACE	1.10	6.14	13.93	11.04	5.75	37.96
All	10.02	16.27	38.22	25.27	10.22	100.00

Table A.3: Summary Statistics of Rating & Duration Portfolios Excess Returns

This table reports the summary statistics of sorted portfolio returns by credit rating and duration length. Panel A reports the summary statistics for five rating-sorted portfolios. Mean (%) and Std (%) denote the time-series average and standard deviation of the portfolio returns. Corr. denotes the time-series correlation between the portfolio returns and S&P 500 index returns. Panel B reports the summary statistics for 25 sorted portfolios by rating and duration. All portfolio statistics are provided with both equal-weighted and value-weighted versions.

Panel A: By Rating

	Equal-weighted			Value-weighted		
Rating	Mean (%)	Std. (%)	Corr.	Mean (%)	Std. (%)	Corr.
AAA	0.08	1.71	0.17	0.15	1.65	0.20
AA	0.13	1.52	0.26	0.16	1.52	0.27
A	0.15	1.58	0.31	0.18	1.65	0.30
BBB	0.23	1.85	0.35	0.23	1.84	0.32
Junk	0.43	2.84	0.45	0.48	2.48	0.45

Panel B: By Rating & Duration

	Equal-weighted		Value-weighted		Equal-weighted		Value-weighted	
Duration	Mean (%)	Std (%)	Mean (%)	Std (%)	Mean (%)	Std (%)	Mean (%)	Std (%)
	AAA				AA			
1	0.02	1.06	-0.04	1.07	0.08	0.95	0.06	0.97
2	0.06	1.50	0.00	1.53	0.11	1.32	0.09	1.30
3	0.22	1.78	0.11	1.80	0.17	1.63	0.14	1.62
4	0.26	1.98	0.15	1.94	0.18	1.90	0.16	1.93
5	0.29	2.51	0.23	2.65	0.29	2.37	0.28	2.44
	A				BBB			
1	0.09	1.12	0.07	1.13	0.14	1.34	0.13	1.41
2	0.13	1.46	0.10	1.46	0.17	1.66	0.17	1.73
3	0.18	1.76	0.13	1.71	0.22	1.94	0.25	2.03
4	0.20	1.96	0.19	1.98	0.24	2.19	0.25	2.12
5	0.28	2.29	0.29	2.41	0.33	2.50	0.35	2.56
	Junk							
1	0.33	2.46	0.40	2.88				
2	0.37	2.69	0.41	3.01				
3	0.48	2.35	0.42	2.77				
4	0.47	2.93	0.36	3.37				
5	0.69	4.20	0.51	4.63				

Table A.4: Robustness Check:
Out-of-Sample Weight (%) for Portfolio Return Combination (unconstrained)

This figure reports the out-of-sample weight (in percentage %) for the portfolio returns combination. The dynamic weight is updated annually with past-five-year data (realized returns) and is used to construct the combination forecast for the current year. No constraints are on the weight sign or summation. The colors of the table cells are conditional on the t-stat. The t-stat is calculated grouping the rating & duration 25 portfolios and each year. Finally, we highlight the table cells by different absolute t-stat values. The dark, medium, light, and blank colors mark significance at the 1%, 5%, 10% level, and no significance. The blank background cells are of minor importance in our combination, whereas the dark background cells are very important.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Rating	-59	-9	-3	26	43	56	60	64	65	68	72	83	67	67	67	66	50	63	69	67
Duration	28	36	40	24	20	19	4	-2	-2	5	6	1	-19	-22	-27	-28	-32	5	3	18
Maturity	16	7	5	11	16	3	33	40	45	42	52	45	26	24	26	26	32	44	51	36
Age	-17	-15	-20	-25	-31	-43	-61	-52	-54	-54	-48	-18	-28	-27	-25	-25	-40	-36	-31	-15
Size	35	35	41	42	47	45	34	26	23	11	8	26	18	19	21	20	13	33	29	27
Short-term Reversal	42	44	42	35	35	53	58	62	65	69	67	43	30	32	35	34	42	63	63	49
Momentum 6 Month	26	21	20	19	-8	-5	-2	-4	-5	3	0	8	10	10	11	11	14	11	16	11
Momentum 12 Month	23	20	23	18	15	31	29	29	31	28	13	22	15	13	13	16	13	9	11	18
Long-term Reversal 2 Year	-1	-6	8	3	6	-2	-7	-7	4	-1	8	1	-10	-11	-10	-12	-9	10	6	2
Long-term Reversal 3 Year	2	-2	1	2	-12	-17	-16	-19	-21	-13	-8	-34	-24	-19	-23	-21	-6	-6	-21	-9
Variance	-44	-44	-45	-23	-15	0	9	13	14	18	18	-2	-2	-5	-6	-9	-1	0	8	-9
Downside Risk	10	6	-4	-9	-9	-2	-3	-3	-5	-9	-11	4	34	36	35	36	35	16	5	15
Skewness	8	12	31	30	39	30	17	5	-4	-27	-45	-5	10	10	10	11	13	-18	-23	-33
Kurtosis	-10	-25	-40	-38	-21	-25	-20	-15	-18	-22	-19	7	-17	-15	-16	-15	-41	-50	-45	-27
Beta_mkt	6	3	-7	2	-6	-19	-26	-22	-26	-23	-18	-16	-11	-11	-9	-8	-4	-9	-4	-2
Beta_smb	14	7	-6	-26	-26	-3	-10	-8	-4	-3	-14	-11	-17	-16	-15	-15	-20	-19	-22	-42
Beta_hml	10	3	6	-6	2	14	16	12	12	16	4	-54	3	2	2	1	18	-26	-26	-30
Beta_term	18	25	33	37	32	12	13	14	14	15	17	2	3	0	-4	-5	3	-1	9	17
Beta_def	-5	-13	-17	-24	-29	-51	-44	-45	-46	-34	-16	16	26	25	26	26	25	3	3	-5
Residual Variance	5	2	3	12	12	16	24	23	23	19	22	-8	-4	0	2	3	10	18	9	18
Absolute t-stat			$[0, 1.645)$					$[1.645, 1.960)$					$[1.960, 2.576)$					$[2.576, \infty)$		

Table A.5: Robustness Check:
Weight (%) for Benchmark Combination
Three-Year Rolling Window Estimation

This figure reports the weight (in percentage %) for portfolio return combinations. The dynamic weight is updated annually with past-three-years data (realized returns) and is used to construct the combination forecast for the current year. The weight sums to unity in each year. The colors of the table cells are conditional on the positive frequency of weights in bootstrap samples. Specifically, we do 1,000 bootstrap sampling for the three-years data. For each bootstrap sample, we refit the weights. Then, we count the positive frequency of each benchmark among the 1,000 repeats. Finally, we categorize the cells into four sets based on the positive frequency, which is divided by three break points—10%, 50%, and 90%—and color the four sets of cells in blank, light, medium, and dark backgrounds. The blank background cells are of minor importance in our combination, whereas the dark background cells are very important. Some cells have 0 weight and a light color, which may be confusing at first glance. The number 0 is inferred from the whole sample; however, in the some bootstrap samples, this predictor may have positive weights. Given the light color, this benchmark is weighted positively in less than 10% of cases; thus, it is not a strong benchmark.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Rating	0	0	0	22	17	36	46	46	40	32	38	72	59	58	38	48	46	40	32	32
Duration	19	18	21	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	12
Maturity	0	0	0	0	9	0	1	0	21	29	30	5	0	0	0	0	6	38	43	42
Age	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Size	13	20	35	39	43	15	0	0	0	0	0	9	0	0	0	0	0	0	0	0
Short-term Reversal	26	28	24	14	13	39	49	49	38	39	33	3	4	5	23	46	49	22	24	14
Momentum 6 Month	17	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Momentum 12 Month	8	0	0	0	0	10	3	0	0	0	0	11	7	7	0	0	0	0	0	0
Long-term Reversal 2 Year	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Long-term Reversal 3 Year	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variance	0	0	0	0	0	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0
Downside Risk	0	0	0	0	0	0	0	0	0	0	0	0	22	25	39	6	0	0	0	0
Skewness	0	0	0	9	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kurtosis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_mkt	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_smb	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_hml	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_term	17	27	20	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Beta_def	0	0	0	0	0	0	0	0	0	0	0	0	8	6	0	0	0	0	0	0
Residual Variance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bootstrap Positive Frequency			[0, 100)					[100, 500)					[500, 900)					[900, 1000)		

Table A.6: Robustness Check:
Weight (%) for Benchmark Combination
Ten-Year Rolling Window Estimation

This figure reports the weight (in percentage %) for portfolio return combinations. The dynamic weight is updated annually with past-10-years data (realized returns) and is used to construct a combination forecast for the current year. The weight sums to unity in each year. The colors of the table cells are conditional on the positive frequency of weights in bootstrap samples. Specifically, we do 1,000 bootstrap sampling for the 10-years data. For each bootstrap sample, we refit the weights. Then, we count the positive frequency of each benchmark among the 1,000 repeats. Finally, we categorize the cells into four sets based on the positive frequency, which is divided by three break points—10%, 50%, and 90%—and color the four sets of cells in blank, light, medium, and dark backgrounds. The blank background cells are of minor importance in our combination, whereas the dark background cells are very important. Some cells have 0 weight and a light color, which may be confusing at a first glance. The number 0 is inferred from the whole sample; however, in some bootstrap samples, this predictor may have positive weights. Given the light color, this benchmark is weighted positively in less than 10% of the cases; thus, it is not a strong benchmark.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
Rating	0	0	0	0	0	15	28	33	36	39	42	58	55	55	55	56	57	56	56	56
Duration	20	22	23	18	15	8	6	3	0	0	0	0	0	0	0	0	0	0	0	0
Maturity	0	0	1	2	4	0	8	11	15	16	14	8	0	0	0	0	0	0	0	0
Age	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Size	15	21	27	30	34	26	11	5	2	0	0	1	0	0	0	0	0	0	0	0
Short-term Reversal	30	33	30	27	28	39	42	43	43	43	43	32	28	28	30	26	22	20	19	17
Momentum 6 Month	16	11	8	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Momentum 12 Month	6	1	1	4	7	11	4	2	0	0	0	1	5	4	2	1	3	5	5	7
Long-term Reversal 2 Year	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Long-term Reversal 3 Year	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Variance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Downside Risk	0	0	0	0	0	0	0	0	0	0	0	0	13	13	13	17	18	19	19	20
Skewness	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Kurtosis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_mkt	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_smb	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_hml	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beta_term	14	11	10	12	12	1	1	3	5	1	0	0	0	0	0	0	0	0	0	0
Beta_def	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Residual Variance	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bootstrap Positive Frequency			[0, 100)					[100, 500)					[500, 900)					[900, 1000)		

Table A.7: Predicting Corporate Bond Sorted Portfolio Returns

This table reports the out-of-sample R^2_{OOS} values (%) of corporate bond sorted-portfolios excess returns. For each corporate bond characteristic, five sorted portfolios are updated monthly. The prediction methods are shown in each column, and the out-of-sample prediction benchmark is the average of each portfolio excess return. For brevity, we don't show all portfolio in this table; instead, we show the out-of-sample R^2_{OOS} of each group of sorting. For example, the first row shows an aggregated out-of-sample R^2_{OOS} for the AAA, AA, A, BBB, and Junk portfolio returns prediction. The statistics are calculated with Equation 21, and the results for benchmarks for rating, duration, age, short-term reversal, and downside risk are also shown in Table 7, noted as "All."

Characteristic	Mean	Lasso	PCA	Random Forest
Rating	2.17	9.19	13.52	3.34
Duration	3.35	16.25	17.43	8.82
Maturity	3.42	17.62	18.30	8.37
Age	3.60	17.76	18.03	11.09
Size	3.61	15.09	15.75	6.87
Short-term Rev	3.50	17.66	16.32	14.05
Momentum 6M	2.85	14.64	14.65	6.84
Momentum 12M	2.76	16.51	12.07	6.72
Long-term Rev 2Y	3.88	16.39	22.03	14.17
Long-term Rev 3Y	4.12	18.86	20.67	13.74
Variance	3.55	10.96	19.52	13.22
Downside Risk	3.14	11.06	16.68	10.59
Skewness	3.98	18.97	16.64	12.80
Kurtosis	4.09	17.08	19.13	13.10
Beta_mkt	3.28	16.59	17.35	9.89
Beta_smb	3.67	13.06	16.80	10.86
Beta_hml	3.44	12.69	16.45	6.63
Beta_term	3.66	17.74	12.92	9.78
Beta_def	3.31	16.35	14.39	11.92
Residual Var.	3.56	16.41	19.16	12.53

Table A.8: Predicting Individual Bond Returns with BCM Forecast, Out-of-Sample R^2_{OOS} %

This table reports the subperiod analysis for out-of-sample R^2_{OOS} values (%) of predicting individual corporate bond excess returns via the BCM forecast. The three subperiods are divided into the 2008 global financial crisis, including 1998-2007 (pre-crisis), 2008-2009 (crisis), and 2010-2017 (post-crisis). In Panel A, the prediction on an individual corporate bond excess return is the BCM forecast. The weights are reported in Table 2. In Panel B, the combination weights are unconstrained, as reported in Table A.4 In Panel C, the predictions are similar to those of Panel A but use equal weight in combination weights. Panel D reports the prediction performance of factor models. The out-of-sample R^2_{OOS} baseline prediction is the belonging credit rating portfolio average excess return. We report the Fama-Macbeth t-statistics, where ***, **, and * indicate significance at the 1%, 5%, and 10% level, respectively.

	1998-2007		2008-2009		2010-2017	
	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$	R^2	FM-$\overline{R^2}$
Panel A: Benchmark Combination Model, Constrained						
c-avg.	0.30	0.49***	0.09	0.21***	0.59	0.93***
c-mean comb.	0.47	0.83***	0.47	0.56***	1.57	1.38***
c-Lasso	1.82	2.44***	3.31	3.30***	2.60	0.49
c-pca	0.76	1.77***	4.34	4.10***	2.30	-0.43
c-rf	2.00	2.56***	0.53	1.44***	3.28	1.90**
Panel B: Benchmark Combination Model, Unconstrained						
u-avg.	0.48	0.67***	0.14	0.29***	1.17	1.74***
u-mean comb.	0.66	1.03***	0.44	0.58***	2.32	2.42***
u-Lasso	2.05	2.59***	2.95	2.84**	2.49	0.89
u-pca	0.66	1.69**	4.25	3.96***	1.85	-0.68
u-rf	2.02	2.49***	0.22	1.18**	3.28	2.27**
Panel C: Benchmark Combination Model, Mean Combination						
m-avg.	0.07	0.14***	0.08	0.11***	0.18	0.30***
m-mean comb.	0.26	0.50***	0.53	0.46***	1.01	0.62***
m-Lasso	1.44	2.01***	3.58	3.00***	2.55	0.36
m-pca	0.74	1.67**	4.25	3.71***	1.99	-0.97
m-rf	1.66	2.14***	1.02	1.33***	3.09	1.50*
Panel D: Factor Model						
FF5	-0.57	-0.75***	-0.08	-0.53	0.12	0.20
BBW4	-0.74	-0.97***	-0.09	-0.56	-0.03	0.10

B Predictor Descriptions

B.1 Corporate Bond Characteristics

B.1.1 Fundamental Characteristics

We have five fundamental characteristics: credit rating (CRT), time-to-maturity (TMT), age (AGE), duration (DUR), and amount outstanding (SIZE).

We collect individual bond level credit ratings (CRT) from Mergent FISD. All ratings are in numerical numbers. For example, 0 refers to a AAA rating, 1 refers to AA+, ..., and 21 refers to D. Investment-grade bonds have ratings from 0 (AAA) to 9 (BBB-). Non-investment-grade bonds have ratings above 9. If both Moody's and S&P ratings are available, we use the average.

TMT is the number of years left for a bond from the current month to the maturity date. AGE is the time between the current trading month and the issuance data, in number of years. DUR is calculated as the Macaulay duration with yield-to-maturity from Gürkaynak et al. (2007). SIZE is collected from Mergent FISD.

B.1.2 Return-based Characteristics

Following Jostova et al. (2013), we have created five past return related characteristics, including STR (lagged one-month returns as short-term reversal), mom6m (lagged two-month to lagged six-month cumulative returns as momentum), mom12m (lagged two-month to lagged 12-month cumulative returns as momentum), ltr2y (lagged 13-month to lagged 24-month cumulative returns as long-term reversal), and ltr3y (lagged 13-month to lagged 36-month cumulative returns as long-term reversal). Following Bai et al. (2019a) and Bai et al. (2019b), we include VAR (variance), DSD (downside risk 5% VaR), SKEW (skewness), and KURT (kurtosis). These four distributional characteristics are calculated using the rolling sample of the past 36 months.

B.1.3 Risk Characteristics

Following Fama and French (1993), we list five important risk factors for corporate bond returns: MKT, SMB, HML, TERM, and DEF. We estimate the multivariate beta of individual bonds on the factors using the rolling sample of the past 36 months. In this part, we have BETA_MKT, BETA_SMB, BETA_HML, BETA_DEF, BETA_TERM, and the corresponding residual variances RVAR.

B.2 Macro Predictors

We have 20 macro predictors in this category: S&P 500 predictors (dividend-to-price, earnings-to-price, net equity issuance, leverage, and stock variance), Pastor-Stambaugh illiquidity, duration spread, rating spread, corporate bond market returns, Fama-French three factors (MktRf, SMB, and HML), term factor, and default factor, one-month Treasury bill rate, inflation, long-term returns, default yield, term spread, and Cochrane-Piazessi forward factor.

B.2.1 Fama-French Three Factors (MktRf, SMB, HML)

We use the Fama-French three factors (MktRf for excess market return, SMB for size factor, and HML for value factor), which are downloaded from French's website[10].

B.2.2 Corporate Bond Market Factor (CBMKT)

Our corporate bond market factor (cbmkt) is the equal-weighted average return of all observed corporate bond excess returns in each month.

B.2.3 Term Factor (TERM)

Following Fama and French (1993), TERM proxies for the unexpected changes in interest rates. It is constructed as the difference between the monthly long-term government bond return (from Ibbotson Associates) and the one-month Treasury bill rate measured at the end of the previous month.

B.2.4 Default Factor (DEF)

Following Fama and French (1993), the default factor arises from the shifts in economic conditions that change the likelihood of default. It is proxied by DEF, the difference between the return

[10] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

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.

B.2.5 3-Month Treasury Bills (TBL)

Following Welch and Goyal (2008), the TBL is the 3-Month Treasury Bill : Secondary Market Rate from the economic research database at the Federal Reserve Bank at St. Louis website[11].

B.2.6 Inflation (INFL)

Following Welch and Goyal (2008), the INFL is the *Consumer Price Index (All Urban Consumers)* from the Bureau of Labor Statistics.

B.2.7 Long-Term Rate of Returns (LTR)

Following Welch and Goyal (2008), the LTR is the long-term rate of returns provided by Ibbotson Associates.

B.2.8 Term Spread (TMS)

Following Welch and Goyal (2008), the TMS is the difference between the long-term yield on government bonds (by Ibbotson Associates) and the Treasury bill. TMS is different from TERM in B.2.3, because TMS uses the long-term yield on government bonds, whereas TERM uses the long-term return on government bonds.

B.2.9 Default Yield Spread (DFY)

Following Welch and Goyal (2008), the DFY is the difference between BAA- and AAA-rated corporate bond yields.

[11] https://fred.stlouisfed.org/series/DTB3

B.2.10 Cochrane-Piazzesi Forward Factor (CP5)

Following Cochrane and Piazzesi (2005), we construct a five-year forward factor with the codes published on John Cochrane's website[12].

B.2.11 Maturity Spread (MTS)

We construct a short maturity portfolio (two to five years) and a long maturity portfolio (longer than 10 years) in each month. The MTS is the spread between the long and short maturity equal-weighted portfolio returns.

B.2.12 Rating Spread (RTS)

The RTS is the spread between the AAA and Junk rating equal-weighted portfolio returns.

B.2.13 Paster-Stambaugh Illiquidity (PSILL)

Following Pástor and Stambaugh (2003), we download the market illiquidity measure from Robert Stambaugh's Website[13].

B.2.14 Equity Market Variables

We download the data in Welch and Goyal (2008) from Amit Goyal's website[14] and calculate the S&P 500 characteristics, including MKTDY (Dividend Yield), MKTEP (Earnings-to-Price), MKTNI (Net Equity Issuance), and MKTSVAR (Stock Variance).

[12] https://faculty.chicagobooth.edu/john.Cochrane/research/Data_and_Programs/Bond_Risk_Premia/index.htm

[13] http://finance.wharton.upenn.edu/~stambaug/liq_data_1962_2018.txt

[14] http://www.hec.unil.ch/agoyal/.