Deep Learning for Corporate Bonds*

Jetlir Duraj
University of Pittsburgh

Oliver Giesecke
Stanford University

March 3, 2025

Abstract

We estimate an asset pricing model for the U.S. corporate bonds market using bond portfolios, as well as a large longitudinal dataset of individual bonds that we augment with fundamental characteristics of the issuer. We further enrich the information set with a large set of macroeconomic time series. We estimate diverse model architectures with two approaches: (1) minimizing the mispricing loss, and (2) maximizing the Sharpe ratio. We find that, contrary to the equivalence of these two approaches in the sense of financial theory, maximizing the Sharpe ratio performs better for individual bonds, whereas the difference is smaller for bond portfolios. The out-of-sample annual SDF portfolio Sharpe ratios are in the range of .59 to 1.00, and show statistically significant excess returns (alphas) relative to conventional risk factors. Our results are robust to the exclusion of financials and REITs.

Keywords: conditional asset pricing, corporate bonds, deep learning, general method of moments, big data, stochastic discount factor, Sharpe ratio
JEL codes: C14, C38, C55, G12, G14

*Duraj: jed169@pitt.edu, corresponding author. Giesecke: ogiesecke@stanford.edu. We thank Hui (Judy) Yue for excellent research assistance in the construction of the dataset. This research was supported in part by the University of Pittsburgh Center for Research Computing, RRID:SCR022735, through the resources provided. Specifically, this work used the H2P cluster, which is supported by NSF award number OAC-2117681. Any opinions, findings and conclusion contained in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. Link to code base for the project.

1 Introduction

Empirical asset pricing for fixed income has traditionally focused on treasury bonds (Cochrane and Piazzesi (2005), Ludvigson and Ng (2009), Bianchi et al. (2021)). U.S. corporate bonds constitute another asset class for fixed income investors. The Securities Industry and Financial Markets Association (SIFMA) estimates a total volume of 10.3 trillion USD outstanding as of Q4 2022. To put this in perspective, the outstanding volume is comparable to the overall mortgage debt in the United States or about half of all debt issued by the U.S. federal government. Despite its size, few studies have examined its return characteristics; some notable exceptions in recent years are Nozawa (2017), Elkamhi et al. (2020), Bali et al. (2021), and He et al. (2021).

This paper estimates asset pricing models for corporate bonds in an environment with a large amount of conditioning information. Hence, the model has to fulfill two tasks: (i) efficient dimensionality reduction and (ii) asset pricing. We use the flexibility offered by neural network architectures to estimate the SDF by modeling the portfolio weights of the SDF on the tradable asset space via neural networks. We contrast these flexible models with conventional approaches that postulate linearity of the SDF in its pricing factors.

In comparison to equity prices and returns, information on corporate bonds have been historically scarce. Our first contribution is the construction of a large longitudinal dataset of individual bond returns and fundamental information of the respective issuer, combining individual bond returns (Elkamhi et al., 2021) with comprehensive bond characteristics, detailed issuer fundamentals, and a large set of macroeconomic time series (McCracken and Ng, 2016), as well as, extended time series as described in Welch and Goyal (2008). In the process, we develop an original algorithm that merges bond returns and information from corporate financial reports (fundamentals). In addition, we build on the work of Elkamhi et al. (2020) and use characteristics sorted bond portfolios. This approach is closer to the approach in the asset pricing literature that works on the level of portfolios (Fama and French, 1993b, 1995, 2012, 2015b) than on the level of individual securities. Only recently, with the advancement in methodology we have seen more studies that directly price individual securities (Chen et al., 2023; Freyberger et al., 2020a; Gu et al., 2020).

For bond portfolios we obtain annual out of sample Sharpe ratios between 0.76 an 0.87 depending on the data environment and model objective. We find that models that incorporate a broad set of macro time series perform better, with both higher Sharpe ratios and lower maximum drawdown. The maximum leverage of the SDF portfolio remains modest at 1.15 for the model with the highest Sharpe ratio. The portfolios show modest but statistically significant excess returns via commonly used risk factors, including the S&P500, Fama French 3 and 5 factor models. Even in terms of predictability, we find explained variation of 64% and cross-sectional R^2 of 88%.

Asset pricing for individual securities, in our case individual U.S. corporate bonds, has traditionally been challenging because of the large idiosyncratic noise and the dimensionality of the data. Our model architectures allows to estimate the weights on the tradable asset space flexibly. Despite these challenges, we obtain annual out of sample Sharpe ratios between .41 and 1.00. Thus, the best model surpasses even the best Sharpe ratio on the bond portfolios. The individual bond data have many advantages over the more aggregated portfolio level data. In particular, it allows for the augmentation of the information set with individual bond level characteristics and a large set of issuer characteristics. This additional information allows for a more targeted asset allocation. In addition, the flexibility of the model to allocate across each individual asset in a flexible way provides greater opportunity for diversification. The model with the highest Sharpe ratio takes on more leverage than in the bond portfolio data. At maximum, the SDF portfolio has a leverage of 4.29. This is also reflected in the higher maximum drawdown which is 29% for the portfolio.

While the Sharpe ratio is associated with a higher volatility, the excess returns with respect to commonly used risk factors are substantial in magnitude and highly statistically significant. The strategy has an annual excess return of 12% vis-a-vis the S&P500 and a t-stat of 3.04. Even against the more demanding Fama-French 5 factor model, the strategy obtains an annual excess return close to 12% and a t-stat of 2.50. The model with the best Sharpe ratio obtains explained variation of 12% and cross-sectional R^2 of 18%. This is comparatively low to other models that minimize the mispricing loss, which achieve explained variation of up to 23% and cross-sectional R^2 of up to 45% but generally show lower Sharpe ratios.

There are several more general insights from the model estimation in diverse data environments and objective functions for training. We find that, despite equivalence in theory, maximizing the Sharpe ratio performs typically better than using the mispricing loss for individual bonds, as shown by the respective out-of-sample Sharpe ratios and compounding annual growth rates (CAGR). For predictability, however, we find that training the model with mispricing loss typically results in higher explained variation and cross-sectional R^2 for individual bonds. Relatedly, we generally find higher and more significant excess returns by maximizing the Sharpe ratio as objective function than minimizing the mispricing loss. This was an unexpected result, as there is ex-ante, no reason to expect a divergence. Intuitively, the mispricing loss is more closely aligned with explaining the variation in returns than maximizing the Sharpe ratio and vice versa. For bond portfolios, Sharpe ratio maximization and mispricing loss deliver comparable out-of-sample Sharpe ratios and excess returns. In terms of predictability, Sharpe ratio maximization performs somewhat better than mispricing loss minimization for bond portfolios.

Another interesting result is that predictability, as measured by explained variation and cross-sectional R^2, always falls without macroeconomic time series. We do not observe this monotonicity for out-of-sample Sharpe ratios, sometimes we even find slightly higher Sharpe ratios when excluding the macroeconomic time series.[1] Interestingly, the exclusion of macroeconomic time series typically results in higher drawdowns, suggesting that macro information is able to predict market downturns and correspondingly low returns. This is complementary to the fact that the exclusion of macro time series often results in higher maximum leverage. This partly, but not entirely, explains the higher volatility and the larger drawdowns when excluding macroeconomic time series.

In comparison to the results of the conventional literature, we contrast our results to a pricing kernel that is linear in yield and macroeconomic factors for 40 bond portfolios estimated via general methods of moments (GMM). The specification mirrors the one of Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) in the treasury bond market. We find that our approach performs better relative to the conventional approach: the Sharpe ratios of the estimated SDF through linear methods are an order of magnitude smaller, and a Diebold-Mariano test comparing the pricing residuals from the SDF estimated via neural networks, with the residuals from predictive regressions returns highly significant negative Diebold-Mariano statistics.

Finally, to complement the results of our neural network approach, we estimated ex-post a hidden Markov model (HMM) on the bond portfolios. The aim is to understand the evolution of regimes for the U.S. corporate bond market with return data independent from the ones estimated through neural networks. We find three distinct well-defined market states: a state with high volatility and relatively high positive returns, a state with low volatility and low but positive returns, as well as a state with high volatility but negative returns. During the time period in our validation and test set, the dominant market regime is one of low volatility and low but positive returns.

[1] This result may not survive a greater hyperparameter and architecture search than the one included in the paper.

Relation to existing literature. The paper connects to four main strands of the literature. First, there is a rich literature on asset pricing in the (treasury) bond market. Fama and Bliss (1987), Cochrane and Piazzesi (2005) and Cochrane and Piazessi (2008) show that forward rates predict excess returns. Cochrane and Piazzesi (2005) and Cochrane and Piazessi (2008) show that a single factor—a convex combination of forward rates—has strong predictive power with an R^2 of up to 0.44 in the treasury bond market. Ludvigson and Ng (2009) further shows that additional macroeconomic factors which are extracted from a large set of macroeconomic time series increases the explanatory power in the treasury bond market marginally. Bianchi et al. (2021) shows that neural networks and extreme trees increase the predictive performance for bond excess returns. The paper finds higher predictive R^2 than Cochrane and Piazzesi (2005) and Ludvigson and Ng (2009) which the authors attribute to the better information processing and the ability to capture important non-linearities by the new set of statistical methods. Our paper replicates the specifications of Cochrane and Piazessi (2008) for corporate bond portfolios and finds R^2s that are much smaller. In contrast to the treasury market, the addition of macroeconomic factors which are derived with a static factor model as in Ludvigson and Ng (2009) increases the explanatory power of expected excess returns substantially. After the inclusion of macroeconomic factors, the R^2 is comparable in magnitude to the full specification of Ludvigson and Ng (2009) in the treasury bond market.

Second, there is a nascent literature that studies asset pricing in the (corporate) bond market. Culp et al. (2018) and Nozawa (2017) construct a micro-dataset on corporate bonds. Nozawa (2017) decomposes the yield spread into expected excess returns and expected credit losses. He finds that the variation in the yield spread originates about half from the volatility of the implied long-run expected credit loss, with the remainder coming from the volatility in the risk premium. Our paper differs from Nozawa (2017) as our focus is not investigating the predictability of excess returns. Elkamhi et al. (2020) tests whether a long-run risk model can account for the observed risk premia on 40 bond portfolios—the test assets. The authors find that a one-factor model based on long-run consumption growth matches the risk premiums on portfolios with moderate risk aversion parameter. Bali et al. (2021) use the implied relationship between expected bond and equity return from a Merton (1974) credit risk model to predict cross-sectional bond returns. He et al. (2021) use a set of machine learning techniques to estimate cross-sectional excess returns, using 20 macroeconomic predictors, 20 corporate bond characteristics, and 20 equity characteristics in the information set. The authors find evidence for out of sample predictability with cross-sectional R^2 of up to 8% for subsets of their sample.

Third, there is a growing literature that uses tools in machine learning for asset pricing. Gu et al. (2020) provides a comparative analysis of machine learning methods for empirical asset pricing. The authors find large (conditional) expected returns conditional on some candidate predictors. The study argues that—in comparison to traditional approaches in asset pricing—neural networks and regression trees can incorporate the vast amount of conditional information and that inter-dependencies among candidate predictors are important. Freyberger et al. (2020b) use an adaptive group LASSO to select firm characteristics that provide incremental explanatory power for expected returns. In addition to the selection of characteristics, the paper uses non-parametric series estimation to allow for non-linearities of expected returns with respect to characteristics but imposes additivity across characteristics to overcome the curse of dimensionality. The paper finds that a small set of characteristics is sufficient to explain differences in cross-sectional expected returns and that non-linearities in the characteristics are generally important. Gu et al. (2021) use an auto-encoder for dimensionality reduction in a conditional asset pricing model. In particular, the paper uses auto-encoder to allow latent factors and factor exposures to depend on conditioning variables. As a result, factor exposures depend in a non-linear way on a small set of hidden variables—the non-linearity of factor exposure differentiates this asset pricing model from the IPCA as presented by Kelly et al. (2019). Bryzgalova et al. (2020) shows how to extract test assets conditional on characteristics via regression trees. The paper shows that the—so extracted—test assets are characterized by higher out of sample Sharpe ratios and expected returns in comparison to state-of-the-art pricing models. Chen et al. (2023) uses the moment-condition as implied by the no-arbitrage to estimate a set of conditional asset pricing models for equities. The paper uses a combination of long-short-term-memory (LSTM) models and feed-forward networks (FFN) for dimensionality reduction and a potentially non-linear approximation of the pricing kernel. Bianchi et al. (2021) shows the success for deep learning for treasury securities. The study documents that macroeconomic time series have increased predictive accuracy beyond predictors that are derived from the term structure for a neural network architecture. This echoes the findings of Ludvigson and Ng (2009) for linear pricing kernels. The authors find evidence for non-linearities among yields, and that macroeconomic time series, play an important role for the pricing of treasury security portfolios. We are contributing to this strain of literature by focusing on U.S. corporate bond returns.

We proceed as follows: Section 2 describes the data construction and sample selection. Section 3 introduces the estimation approach and the evaluation methodology. Section 4 presents the results for each method of estimation and data set. Section 5 compares the results against the literature and other model architectures. Section 6 concludes. The appendix contains various auxiliary technical results that are not part of the main body of the paper, a discussion of alternative estimation architectures, the results of predictive regressions for the linear models, and finally, summary statistics.

2 Data

We use two main dataset for the estimation of asset pricing models for corporate bonds. First, we use a dataset of 40 bond portfolios. Second, we build a comprehensive longitudinal dataset, including a rich set of conditioning information, for individual corporate bonds. While the data for equity and treasury securities is, to a large extent, standardized and readily available, the data for individual corporate bond returns requires several pre-processing steps. In addition, we develop an original algorithm to match bond returns with the fundamental characteristics of the issuer, including the historical equity performance and characteristics. Section 2.1 provides details on the data construction and Section 2.2 describes the sample selection, tabulates summary statistics for each sample, and illustrates the sample composition in comparison to a commonly used benchmark portfolio, the S&P500. Section 2.4 elaborates on the data management that allows for an efficient estimation with limited hardware requirements for a large, unbalanced longitudinal dataset.

2.1 Individual bond data construction

Bond Returns The data construction on corporate bond returns is based on Elkamhi et al. (2020).[2] The data consists of a panel of clean corporate bond prices between February 1973 and December 2019 and is constructed from four main sources, that is, the Lehman Brothers Fixed Income Database, the Mergent FISD/NAIC Database, TRACE, and DataStream. If price information are available from more than one source, priority is given according to following order: Lehman Brothers Fixed Income Database, TRACE,

[2]We thank Yoshio Nozawa for generously sharing his dataset with us.

Mergent FISD/NAIC and DataStream.[3] Bond rating information are obtained from Standard & Poor's when available, and Moody's otherwise. Further, bond defaults are identified from Moody's Default and Recovery database which provides a historical record of bond defaults and recovery values from 1970 onward. If the price information for a defaulted bond is unavailable, we resort to price information in Moody's recovery file to accurately account for the return upon default.

We follow Elkamhi et al. (2020) and apply several filters to the dataset. First, we exclude bonds with floating rates and other embedded options other than callable bonds. While call-ability provides optionality and is typically priced, Chordia et al. (2017) shows that the call-ability of corporate bonds does not affect the cross-section of average returns in a significant way. In addition, excluding call-able bonds would greatly reduce the length of the sample period because–until the late 1980s–very few bonds were non-callable. Furthermore, we exclude a few observations that are likely to erroneous recording, that is, we exclude price observations that are below 5 dollars or above 1,000 dollars per 100 dollar face value. Further, we remove bonds maturing in less than a year.

Based on clean prices and accrued interest, we calculate monthly bond returns as:

$$R_{t+1} = \frac{P_{t+1} + AI_{t+1} + Coupon_{t+1}}{P_t + AI_t} - 1$$

where P_{t+1} is the clean price and AI_{t+1} is the accrued interest at the end of month $t+1$, and $Coupon_{t+1}$ is the coupon paid during month $t+1$.

The data spans a total of 38,955 unique securities, 7,995 unique issuers and 2,341,941 non-missing monthly return observations. The full set of portfolios are available for the period between February 1973 and September 2020.

Macroeconomic time series The majority of the monthly time series data is obtained from the FRED-MD database as constructed and described in McCracken and Ng (2016) and available at `https://research.stlouisfed.org/econ/mccracken/fred-databases/`. All variables are transformed to obtain stationary series according to the recommendations of McCracken and Ng (2016).

We append the macroeconomic time series data with six additional time series from Welch and Goyal (2008).[4] Instead of using the most recent published data by Welch and Goyal (2008) which covered the period until December 2019 at the time of our dataset construction, we reconstruct all time series from conventional data sources which include WRDS, FRED and Bloomberg. The reconstruction allows us to cover the full time period from February 1973 to September 2020 for which bond returns are available. Additional details on the series construction can be found in Appendix Table A.23. Summary Statistics are tabulates in Table A.11, A.12, and A.13.

We exclude the macro variables ACOGNO (Manufacturers' New Orders: Consumer Goods) and UMCSENT (University of Michigan: Consumer Sentiment) as this would shorten our sample period.

Bond issuance and issuer matching While the link between equity prices / returns and fundamentals is well established via the Compustat/CRSP crosswalk, the link between bond issuance and issuer is generally missing. We construct such link based on our own matching algorithm. Each bond issuance is uniquely identified by the 9 digit CUSIP, where the first 6 digits identify the issuer, the following 2 digits the security

[3] Nozawa (2017) shows robustness to reversing the order of priority in the online appendix.

[4] This includes (i) dividend-price ratio, (ii) book-to-market ratio, (iii) net equity expansion, (iv) term spread, (v) default spread, (vi) stock variance.

and the last digit is the check number. Importantly, the issuer identifying information for debt instruments differs from those of equity instruments; hence, a naive matching on the first 6 digits leads to erroneous results. Instead, we use a combination of Mergent FISD, the S&P rating file, and the Capital IQ Identity file to obtain information about the issuer of the debt security. Ultimately, we declare a match to be valid if the corresponding gvkey $\times$ year combination has non-empty or non-zero asset in Compustat. If more than one valid gvkey is identified, we resolve any ambiguity based on an iterative name matching using the match that exhibits the lowest Levenshtein distance. We conducted extensive validation and found good matching results. The overall matching rate using this procedure at the year $\times$ CUSIP level was 78.7% . We conducted a study of the unmatched bond issuances and found a few patterns. First, public utility companies in the U.S. are active bond issuers but often remain private companies which are not covered by Compustat. Further, we found a few issuances where the Compustat coverage seemed incomplete for unknown reasons. We found no evident pattern that would introduce a selection bias apart from a selection on public companies as dictated by the nature of Compustat.

Issuer fundamentals and equity characteristics With the cross-walk between bond returns and Compustat at hand, we construct a comprehensive dataset on issuer fundamentals based on balance sheet characteristics in Compustat. This includes, among others, total assets, cash position, cash flow information and profitability indicators. Further, we use the established cross-walk between Compustat and CRSP to append further market based information; such as, share price, shares outstanding and derived variables as market capitalization and market-to-book ratios. In addition, we use CAPM and Fama-French 3-factor estimates on factor exposure, alpha, and idiosyncratic volatility from the WRDS Beta Suite. Further, we add lagged equity returns for bond issuers that have publicly traded equity outstanding. We follow Daniel et al. (2020) to adjust the equity returns for delistings.

The full list and the description of fundamental variables and market based variables is tabulated in Appendix Table A.20 to Table A.23.

Bond portfolios We use the bond portfolios constructed by Elkamhi et al. (2020). The portfolios are sorted based on the following characteristics, that is, (i) credit spreads (Nozawa (2017)) (ii) intermediary capital exposure (He et al. (2017)), (iii) credit rating (Bai et al. (2019)), (iv) downside risk (Bai et al. (2019)), (v) maturity (Gebhardt et al. (2005)), (vi) idiosyncratic volatility (Chung et al. (2019)), (vii) and long-term reversals (Bali et al. (2021)). All portfolios are value-weighted. For credit spreads we create a portfolio for each tercile and a portfolio for each quintile for the other characteristics.[5] The portfolio returns cover the period from February 1977 to December 2018.

2.2 Summary statistics

We obtain two final dataset, one for bond portfolios and the other for individual bonds. In the following we provide a summary for each of the dataset, more detailed summary statistics are in the Appendix.

Bond portfolios We enrich the dataset with the 40 bond portfolio returns with a broad set of macroeconomic time series, Goyal & Welch time series, the cross-sectional median of corporate fundamentals, and

[5] Portfolio returns between period t and $t+1$ are obtained by sorting bonds based on their characteristics in the time period between month $t-12$ and $t-1$. The sorting and the measurement of the return is done with a gap of one month to minimize the impact of measurement errors in bond prices. Long-term reversals are measured by the negative of cumulative 3-year returns from $t-48$ to $t-12$.

bond portfolio specific return characteristics, such as the spread over treasures, momentum, reversal amongst others. Table 1 summarizes the sample and Appendix Tables A.8, A.9 tabulate selected summary statistics.

	value
# of macro variables	126
# of Goyal-Welch Variables	6
# of median time series fundamentals	41
# of forward spreads	10
# return characteristics	9
# of portfolios	40
Sample start	October 1978
Sample end	December 2018

Table 1: Overview Bond Portfolio Dataset

Individual bond returns The dataset on individual bond returns comprises individual bond returns, macro time series, and corporate fundamentals, lagged equity returns and factor exposure. Table 2 summarizes the sample and Appendix Table A.10 tabulates selected summary statistics.

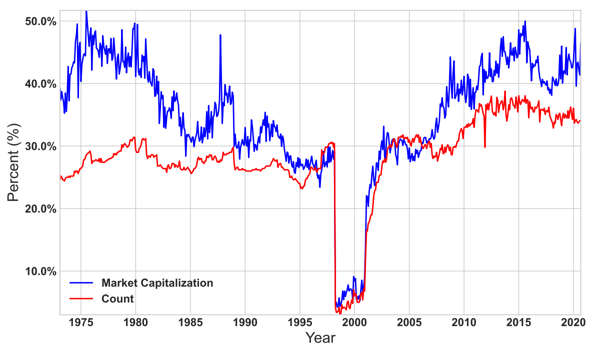
	value
# of firm fundamentals	41
# of macro variables	126
# of Goyal-Welch Variables	6
# of industry dummies	10
# of bond characteristics	9
# of forward spreads	10
# of median time series fundamentals	41
# securities	10,715
# issuers	1,922
Sample start	February 1973
Sample end	January 2020

Table 2: Overview Individual Bond Return Dataset

2.3 Sample representativeness

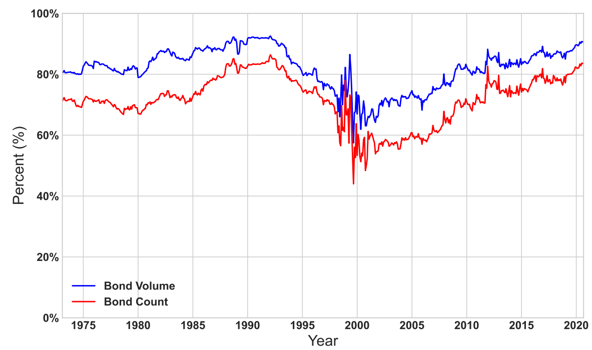
The data collection and selection raises the question of which firms account for the majority of our sample. We illustrate the sample against one common benchmark, the S&P500. Figure 1a shows the share for which the issuers in our sample account for in the S&P500. With the exception of the 1998-2001 period, our sample accounts for 30% or more of the market capitalization, and approximately 30% of the number of firms in the S&P500. The 1998-2001 period constitutes an outlier because the coverage of the Lehman Brother Fixed Income Database fades out but the TRACE trade reporting system was not yet established. There is no alternative data source for this time period and thus this limitation of the data seems insurmountable.

We illustrate the fraction of our sample bonds that are part of the S&P 500. Figure 1b shows that over the sample period, between 60%-80% of all bonds in our sample are issued by firms which are part of the S&P 500; volume wise the percentage is approximately 10% higher. The small decline around the year 2000 coincides with the time period in which the number of public firms reached its historical peak in the U.S. As a consequence, it is more likely that our sample contains public firms that do not belong to the S&P500. While the share shows some volatility in the 1998-2001 period, the share and volume of the covered sample as a share of the S&P500 remains approximately constant. In untabulated summary statistics, we see that this is primarily a result of the shrinking cross-section, not so much of the sample composition. Thus, while the 1998-2001 period remains a data limitation, the representation of firms during this period is somewhat stable.



(a) Sample as Share of S&P 500

Notes: The figure shows the number of companies as a share of the total number of companies in the S&P 500 (Count) and the market capitalization of the companies in the sample as of the total market capitalization of the S&P 500 (Market Capitalization).



(b) Fraction of the Sample in S&P 500

Notes: The figure shows the fraction of bonds in our sample that are in the S&P 500 (Bond Count) and the fraction of the face value of the bonds in the S&P 500 (Bond Volume).

Figure 1: Sample Composition

2.4 Data management

One of the challenges with financial data is a typically unbalanced longitudinal dataset regardless whether one studies equities or bonds. Firms enter and exit public markets which results in a lot of turnover for datasets with a long sample horizon. The problem is only exacerbated with bonds as most bonds have a varying maturity between 3 to 7 years. The unbalanced nature of the data introduces serious challenges in the estimation. One solution in the literature is to expand the sample with unknown values to create a balanced longitudinal dataset (see e.g. Chen et al. (2023)). While this solves the technical problems at estimation, it introduces other challenges. First, it introduces a look-ahead bias as the sample universe in the future is not known at the present. Second, it significantly increases the hardware requirements and the need for parallelization due to memory constraints. Parallelization in turn makes gradients more noisy.

Our approach is to augment the set of features we feed to the algorithm with a double index, that uniquely identifies an asset over time. When a batch of data is furnished to the algorithm, it first splits the batch according to the dates present in the batch, and then processes the dates sequentially. This gives the algorithm only the minimal information it needs to calculate the weights of the assets, and the loss for the current data batch. As such, we are able to estimate all of our models on a single A100 GPU with 40GB memory. Moreover, this allows to use the batch size as a varying hyperparameter for training the neural network architectures. Batch size is an important hyperparameter for all stochastic gradient descent optimization algorithms.

3 Methodology

We estimate the stochastic discount factor (SDF) that prices corporate bonds in the presence of a large set of macroeconomic and individual fundamentals as conditioning information. We first discuss the theoretical foundations for the estimation via mispricing loss in Section 3.1. We then introduce our two main approaches for the estimation of the model architectures in Section 3.2. Section 3.3 introduces the model candidates and other important estimation details; including, the regularization of the estimation. Last but not least, we discuss the evaluation of our models in Section 3.4.

3.1 No arbitrage and mispricing loss

We derive the mispricing loss from the conditional moment restriction that is at the heart of finance theory. The conditional moment restriction is a natural starting point since it is implied by no arbitrage (Cochrane, 2009). Concretely, no arbitrage implies a conditional moment restriction of the following form:

$$\mathbb{E}_t[M_{t+1}R^e_{t+1}] = 0, \tag{1}$$

where M_{t+1} is the stochastic discount factor, R^e_{t+1} is the excess return of the asset over the risk-free rate, and $\mathbb{E}_t$ is the expectation with respect to the time t information set.

We consider a linear SDF with a single risk factor of the following form $M_{t+1} = 1 - \omega_t^T R^e_{t+1} = 1 - F_{t+1}$. Using this formulation, we can express the excess return as:

$$\mathbb{E}_t[R^e_{t+1}] = \left(-\frac{Cov_t(M_{t+1}, R^e_{t+1})}{Var_t[M_{t+1}]}\right)\frac{Var_t[M_{t+1}]}{\mathbb{E}_t[M_{t+1}]} =: \beta_t\lambda_t. \tag{2}$$

We can now reinterpret equation (2) as a regression written as:

$$R^e_{t+1} = \beta_t\lambda_t + \epsilon_{t+1}. \tag{3}$$

Thus, we can express the mispricing error as:

$$\hat{\epsilon}_{t+1} = R^e_{t+1} - \hat{R}^e_{t+1} = \left(\mathbf{Id} - \hat{\beta}_t(\hat{\beta}_t^T\hat{\beta}_t)^{-1}\hat{\beta}_t^T\right)R^e_{t+1}. \tag{4}$$

Our approach will be to estimate the weights ω_t that constitute the risk factor $F_{t+1} = \omega_t \cdot R^e_{t+1}$ using various features and neural network architectures, where each time the loss is guided by minimizing the empirical deviation from the fundamental asset pricing equation in (1). Written out fully, the empirical loss we use is given by:

$$\frac{1}{N}\sum_{i=1}^{N}\frac{|T_i|}{T}\left\|\frac{1}{|T_i|}\sum_{t\in T_i}\left(1-\omega(z_{t,\cdot})\cdot R^e_{t+1}\right)R^e_{t+1,i}\right\|^2. \tag{5}$$

where $\|\cdot\|$ is an euclidean vector norm and $M_{t+1} = 1 - \omega(z_{t,\cdot}) \cdot R^e_{t+1}$ is the estimated SDF. A more general form would include a vector of test assets $g(\cdot)$, and is discussed in section 5.2. An important implication of this approach is that all the estimated SDFs are on the space of tradeable assets and thus can be traded as a strategy. The resulting asset pricing model is a conditional asset pricing model, since all the portfolio model weights ω_t are conditional on the current information set.

Once the SDF is estimated, the definition of the β_t in equation (2) can be used to estimate the risk exposures of the assets. Based on the identity $E_t[M_{t+1}] = 1$, which can be shown to follow from the definition of the SDF (see e.g. pg. 54 in Back (2017)), we estimate β by estimating

$$E_t[F_{t+1} R^e_{t+1}]. \tag{6}$$

We approach this estimation as a traditional prediction problem with neural networks, where the feature set is the set of available information at time t, and the labels are given by $\hat{F}_{t+1} R^e_{t+1}$, with $\hat{F}_{t+1}$ being the estimated tradeable risk factor of the SDF.

3.2 Estimation approach

For the estimation of our model architectures we use two approaches. We already introduced the conceptual foundation for the minimization of the mispricing loss in Section 3.1. For the second approach, we make use of another theoretical property of the SDF: the SDF projected into the asset space is the portfolio that obtains the highest Sharpe ratio among all tradable portfolios. This leads to our second approach in which we maximize the Sharpe ratio of the SDF portfolio. Thus, we use the following two approaches as the objective function for our model estimation:

A. **Approach 1 - Minimizing mispricing-loss (MP-min):** The SDF is defined by the conditional moment restriction implied by no-arbitrage. We consider the deviation from this moment condition as the mispricing. As such we can define a loss over all mispricing errors which becomes the objective in our estimation procedure.

B. **Approach 2 - Maximizing Sharpe ratio (SR-max):** Using the property that the projected SDF into the space of traded assets exhibits the highest Sharpe ratio, we can use the Sharpe ratio of the SDF portfolio as an alternative optimization objective.

Classic finance theory predicts that both approaches should result in the same portfolio. Foreshadowing some of the result, in our exercise with bond level data, using the maximization of the Sharpe-ratio as training objective, delivers typically higher out-of-sample Sharpe ratios than using the mispricing loss as training objective. For predictability, as measured by explained variation and cross-sectional R^2, we obtain better results using the minimization of the mispricing loss as the training objective. The discrepancy between the two approaches is smaller for bond portfolios.

Another choice that we have to make is the universe of conditioning information that we use in the estimation process. For the estimation at the bond portfolio level we have a set of bond portfolio return characteristics and an exhausting list of macro time series available. When we estimate the SDF at the individual bond level, we have in addition individual bond characteristics and a large set of fundamental characteristics of the respective issuer available. We consider three main data environments in the estimation procedure:

A. Estimate SDF with the full feature universe (baseline).

B. Estimate SDF using only bond characteristics, no macro data (noMacro)

C. Estimate SDF on non-financials and non-REITs only (nonFin)[6]

[6] Because the size of fins and REITs is small, and even smaller in train and validation dataset, we do not perform a full search across all architectures and hyper-parameters. Instead, we re-estimate the best architecture and using the best hyperparameters of the baseline with the slightly smaller nonFin dataset.

Our estimation approach has a set of shortcomings that may affect performance and applicability for investment strategies. One property of our estimation strategy is that it does not require to estimate covariances among assets. While this is not required for the estimation of the SDF and thus greatly simplifies the estimation, we acknowledge that it complicates typical risk management procedures for tradeable strategies, which include the management of draw-down and risk budget periodically. Moreover, our models do not account for trading costs. We acknowledge that trading cost are not negligible in the corporate bond market and may affect the performance of our results, more so than e.g. for equities or ETFs.

3.3 Architectures, hyper-parameters and other details of estimation

Sample Splitting We divide our sample into training, validation, and test set. Specifically, for individual corporate bonds we have a total of 560 dates, of which we use the first 236 dates as part of our training set, 156 for the validation set, and the remaining 168 dates for the test set. The sample for bond portfolios is slightly shorter with a total of 437 dates. We use a similar proportional sample division, with the first 183 dates in the training set, the next 122 in the validation set, and the remaining 132 dates in the test set. For individual bonds this results in a test set comprising the period February 2006 to January 2020. For bond portfolios, the test set comprises the period from August 2008 to December 2018.

Objective function for training. In the SDF estimation, we distinguish between the two approaches. In the mispricing loss approach, we use as a training objective function the mispricing loss 5. The main idea is to choose the weights of the SDF in a way that the total mispricing loss of all observable assets is minimized.

In the Sharpe ratio maximization approach, we maximize the Sharpe ratio of the implied SDF portfolio. For training purposes, the fractional definition of the Sharpe ratio shows some undesirable properties. Since the standard deviation is in the denominator, SGD-based algorithms tend to pick an asset allocation that minimizes the variance of the portfolio with very large or infinite Sharpe ratios during training. This leads to a model for weights that does not generalize out-of-sample. Instead we use a linear version of the Sharpe ratio during training, but use the original definition in the validation and test set. The linear version of the Sharpe ratio still balances the first and the second moment of the return distribution, but does not lead to unrealistic Sharpe ratios in the training set. Concretely, we use the following linearized Sharpe ratio:

$$SR\left(R^e_{t+1}, \omega(z_{t,\cdot})\right) = \text{t-series-mean}\left(\omega(z_{t,\cdot}) \cdot R^e_{t+1}\right) - \gamma \times \text{t-series-stdev}\left(\omega(z_{t,\cdot}) \cdot R^e_{t+1}\right). \tag{7}$$

where γ can be interpreted as a parameter of risk-aversion, that controls the relative importance of the mean and the standard deviation of the distribution.[7] For all of our estimations, we work with a constant risk aversion parameter of 0.01.

One advantage of this approach is that it does not require the estimation of the full variance co-variance matrix, which can be challenging. Another advantage is that we are directly optimizing over one of the metrics of interest, thus bypassing the potential misalignment between the optimization and economic objective. A disadvantage is that performance is evaluated on a batch-level, whereas a production-level model would require performance to be evaluated in a more online fashion.

For beta-network estimation, we use mean-squared error as a training criterion. This, since the estimation is formulated as a standard predictive regression problem with neural networks.

[7] A similar training function appears in Guijarro-Ordonez et al. (2021).

Selection criteria for architecture and hyperparameters In the SDF estimation, for both approaches we use the typical fractional definition of Sharpe ratio to select best hyperparameters and architectures. In the beta-network estimation, we use mean squared error as our loss function.

Model Architectures We use a set of four candidate architectures. We use these to model both the portfolio weights, and the betas as a function of bond and macro characteristics.

A. **Model 1 - Shallow FFN:** hidden units = [64,32,16,8]

B. **Model 2 - Deep FFN:** hidden units = [64,64,32,32,16,16,8,8]

C. **Model 3 - Deeper FFN:** hidden units = [64,64,64,32,32,32,16,16,16,8,8,8,4]

D. **Model 4 - Elastic Net:** hidden units = [1].

We find that each of the aforementioned architectures is chosen as a best architecture based on validation results, for some data environment and objective function. Thus, none of the listed architectures is redundant.

Activation Function There is a growing literature on the selection of the activation function. For our application, we chose SiLU as activation function. SiLU is a smooth approximation of ReLU. While it has been found to be slower in training than ReLU or LeakyReLU, it produces better gradients which do not suffer from the indeterminacy at the kink, like ReLU does (see e.g. Hendrycks and Gimpel (2016)).

Optimizer We use the Adam optimizer (Kingma and Ba (2015)), an optimizer that accelerates and stabilizes gradient descent by keeping track of the exponentially weighted averages of the gradients and squared gradients. It incorporates in its gradient updates information about the curvature of the loss function, as well as momentum. This allows it to be more robust in environments where non-stationary data are present, like ours. Adam optimizer has shown excellent performance in a wide set of applications. The learning rate is a hyper-parameter over which we loop during the training and validation phase. We consider learning rates of 0.0001, 0.001, and 0.005.

Batch size Batch size is an important hyperparameter for every SGD-based optimization algorithm. We try batch sizes of 256, 512, 1024, and 2048.

Number of epochs For both SDF and beta estimation, we train and validate for a maximum of 50 epochs.

Regularization Regularization in the training phase is important to achieve generalization of the model out of sample. It becomes even more important in an environment with low signal-to-noise ratios as ours. We combine several methods for generalization. First, we always **standard-scale features** before feeding them to the neural networks architecture. We further use **L^1-regularization and L^2-regularization** with varying hyper-parameters in the training stage. Specifically, we try ["none", 0.001, 0.01, 0.1] for L^1- and L^2-regularization. Further, we use **batch-normalization** for all architectures, except for the ENet. We also use **early stopping** with a stop after 8 periods of non-improvement in the validation phase.

Lastly, we use a fixed **dropout** of 0.01. We do not experiment with a higher degree of dropout due to known issues of interaction of dropout and batch normalization (see e.g. Li et al. (2019)).

Both model space and hyper-parameter space create a large candidate space of 768 possible combinations that we search over in the validation phase to decide on the best architecture and hyperparameters.

Additional regularization based on leverage Financial returns data typically exhibit a low signal-to-noise ratio which makes generalization of models trained through machine learning methods particularly challenging. In our context, we have found that the weights produced by neural networks do not generalize well without an additional normalization. It is appropriate in such context to employ additional regularization steps that are based on domain knowledge.

One important economic metric in our context, is the amount of leverage that is embedded in the portfolio. Without constraints, the portfolio weights can assume large positive and negative values. As leverage is directly tied to portfolio weights, we normalize the weights produced by the networks via an Euclidean norm. This significantly improves performance and leads to robust results.

There are two Euclidean norms one can possibly use: L^1-norm $\|\cdot\|_1$ and the L^2-norm $\|\cdot\|_2$. The L^1-norm relates directly to the amount of leverage of a portfolio. L^1-norm typically implies much more concentrated portfolios, with some larger weights and lots of smaller weights, L^2-norm typically implies more balanced portfolios with no assets dominating the portfolio. On the other hand, L^1-norm typically implies smaller trading activity focused on a small number of assets, than the L^2-norm. Hence, especially when working with a large cross section that may allow diversification effects, there is a trade-off between diversification and trading costs. As we do not explicitly model trading costs in the paper, we have focused on the L^2-norm. We perform the normalization on the batch level, by putting an upper bound on the Euclidean norm of the weights vector on the batch level.[8] Because of the norm inequalities for Euclidean norms, for a batch of size n there is the following bound between the two norms for the same set of weights ω:

$$\|\omega\|_1 \leq \sqrt{n}\|\omega\|_2.$$

Hence, a bound of C on the L^2-norm can translate to a varying amount of leverage of the SDF on the date level, depending on how many dates are present in the batch. This additional flexibility in endogenous leverage choice, is another advantage of using the L^2-norm to the L^1-norm.

Heuristically, if the typical date of the asset space contains m assets, the number of dates in a batch of size n is typically $\frac{n}{m}$. This translates into a heuristic upper bound on leverage on the date-level of

$$\frac{Cm}{\sqrt{n}},$$

for a batch of size n.

Concretely, we use an upper bound for the L^2-norm of $C_{port} = 1$ for bond portfolios. Depending on the batch size, this practically leads to an effective allowed typical leverage in the range of 0.8 to 2.5. For individual bonds, the number of assets per date in the cross-section has a lot of variability, but is orders of magnitude larger than for bond portfolios. Hence, we choose a smaller bound for the L^2-norm of $C_{bonds} = 0.1$. Depending on the batch size, this choice practically leads to an effective allowed typical leverage range between 2 and 7.

Ensembling We ensemble the winning architecture and hyper-parameters with 6 different random states. All reported results in the paper are of the ensembled models.

Robustness For the nonFin data environment, we do not redo the architecture / hyper-paramater search, but instead re-estimate the winning hyperparameters and architecture. This is because there are relatively

[8] As a byproduct, the effective leverage of the portfolio on the date level is influenced by the batch size used for training.

few fins and REITs in the dataset: only 403 out of 10,715 cusips correspond to financials and REITs overall, and their distribution is skewed towards the test set. For the noMacro data environment we re-estimate everything from scratch.

3.4 Model evaluation

We evaluate all final models based on economically motivated financial metrics as commonly used in the literature. The main interest is in the performance of the model out of sample. We use the Sharpe ratio, but also the commonly used Sortino and Calmar ratios as an evaluation criterion. We also measure other financial variables of interest for the portfolio produced by the model on test set, such as CAGR (cumulative annual growth rate), maximum drawdown (Max-DD in the tables below), but also mean-, and max-leverage (Mean-lev, and Max-lev in the tables below).

Another important benchmark is the model performance in relationship to frequently used benchmarks in the financial literature, or risk factors. Thus, we evaluate the model performance against the the S&P 500 index, the Fama-French 3 factor model (FF3) (Fama and French (1993a)) as well as the more demanding Fama-French 5 factor (FF5) model (Fama and French (2015a)). These are commonly used benchmarks in the asset pricing literature for equity and are commonly accepted risk factors to which most of the excess returns can be attributed. While our assets constitute of corporate bonds, it is reasonable that, at least in part, the exposure to these risk factors also drives excess returns for corporate bond assets.

A separate set of metrics measure out-of-sample predictability. Since predictability at the single asset level is very low (see table of an xgboost exercise in section A.7 of the appendix, which shows negative oos-R2 for bond-level data), we focus on average predictability over the cross-section, or over time. This follows Chen et al. (2023), section F. Concretely, we use the explained variation (EV) and the cross-sectional R^2 (XS-R^2). The explained variation is the R^2 of a cross-sectional regression on the loadings; this is a time-series R^2 and defined as the following:

$$EV = 1 - \frac{\frac{1}{T}\sum_{t=1}^{T} \text{ cross-sec-mean } \left(\epsilon_t^2\right)}{\frac{1}{T}\sum_{t=1}^{T} \text{ cross-sec-mean } \left((R_t^e)^2\right)}. \tag{8}$$

The cross-sectional R^2 is the average pricing error, divided by the average return; this is a cross-sectional R^2, and defined as the following:

$$\text{XS-}R^2 = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N} \frac{T_i}{T} \left(\text{t-series-mean } (\epsilon_i)\right)^2}{\frac{1}{N}\sum_{i=1}^{N} \frac{T_i}{T} \left(\text{t-series-mean } (R_i)\right)^2}. \tag{9}$$

4 Results

We present results at the bond portfolio level, as well as individual bond level. The bond portfolios have the advantage of averaging out idiosyncratic variation and are closest to most asset pricing studies for equity. One of the downsides is that we cannot include issuer fundamentals as conditioning information and the model does not have the flexibility to choose the allocation within each of the bond portfolios. Thus, we estimate the models also at the individual bond level. This takes full advantage of the issuer fundamentals and allows for a more flexible asset allocation. Ex-ante it is unclear which of the approaches yield better results due to a variety of trade-offs. We first discuss the results at the bond portfolio level and then the results at the individual bond level. As outlined in Section 3.2, we estimate each model architecture and hyper-parameter configuration with two separate objective functions.

4.1 Bond portfolio results

We first present the results for 40 characteristics sorted bond portfolios (Elkamhi et al., 2020). As described in Section 3, we estimate asset pricing models for the bond portfolios with two separate sets of conditioning information. First, we use bond portfolio characteristics and a broad set of macro time series. In a second estimation, we omit the macro time series and estimate the models based on the individual bond characteristics only.

For bond portfolios, we find that the best model has a monthly Sharpe ratio of 0.25 (annual Sharpe ratio of 0.87). We find little difference in Sharpe ratios using the Sharpe ratio maximization (Approach 1) vis-a-vis the mispricing-loss minimization (Approach 2) as shown in Table 3. The omission of macro time series leads to overall worse performance. In particular, the best model achieves a lower Sharpe ratio of 0.22, Sortino ratio, and lower statistical significance for the excess returns. In addition, we see higher maximum drawdowns, and higher leverage, relative to the model that uses macro time series as conditioning information.

Environment		**Results Test Set**						
Objective	Data	Sharpe	Sortino	Calmar	Max-DD	CAGR	Max- lev	Mean-lev
MP	Bondpf + Macro	0.25	14.09	0.03	0.06	0.02	1.15	0.44
SR	Bondpf + Macro	0.24	12.28	0.03	0.07	0.02	1.14	0.44
MP	Bondpf	0.22	5.10	0.03	0.17	0.05	2.70	0.99
SR	Bondpf	0.22	12.04	0.03	0.09	0.03	1.23	0.56

Table 3: Bond Portfolios - Performance

Notes: The environment describes the objective function that was used for training and the conditioning information. The best model for each environment is tabulated in Table A.1 in the appendix. The results show a set of financial ratios as described in more detail in Section 3.4, as well as maximum drawdown (Max-DD), and the portfolio implied maximum leverage (Max-lev) and mean leverage (Mean-lev). Sharpe, Sortino and Calmar ratios are monthly. All units are expressed in decimal points.

Figure 2 shows the cumulative returns of the SDF portfolio on the test set for both methods and a full set of characteristics (both bond portfolio characteristics and macro time series).

Excess returns We test for excess returns against commonly used benchmarks, specifically we benchmark against the S&P 500, the Fama-French 3 factor model (Fama and French, 1993b) and the Fama-French 5 factor (FF5) model (Fama and French, 2015b). Table 4 shows the excess returns (alphas) for bond portfolios. For bond portfolios we find statistically significant excess returns of the constructed SDF portfolio vis-a-vis the FF3 factors, marginally significant excess returns against FF5 and robustly significant returns against the S&P500 when we estimate the model with the mispricing loss as the objective function. The out-performance is of magnitude 0.1% at a monthly basis, or 1.2% annually. While the magnitude is similar, the statistical significance is marginally reduced when we estimate the model with the linearized Sharpe ratio as the objective function. We obtain slightly weaker results when omitting the macro time series as conditioning information from the model.

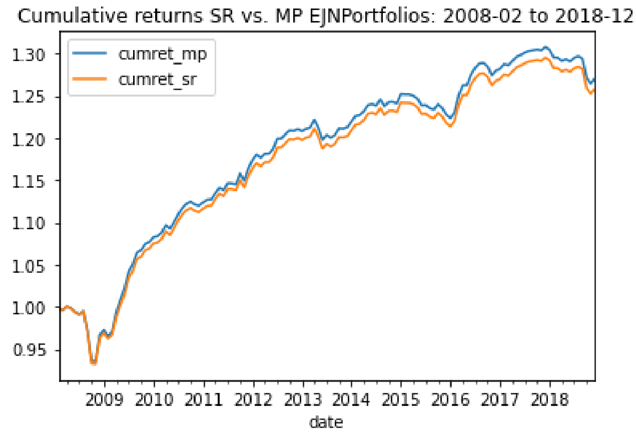


Figure 2: Bond Portfolios - cumulative returns

Environment		**Alphas in Test Set**		
Objective	Data	S&P 500	FF 3 Factors	FF 5 Factors
MP	Bondpf + Macro	**0.001**	**0.001**	0.001
		(2.220)	(2.070)	(1.850)
SR	Bondpf + Macro	**0.001**	0.001	0.001
		(2.100)	(1.900)	(1.700)
MP	Bondpf	0.003	0.003	0.002
		(1.920)	(1.770)	(1.570)
SR	Bondpf	0.002	0.001	0.001
		(1.900)	(1.800)	(1.670)

Table 4: Bond Portfolios - Excess Returns

Notes: The environment describes the objective function that was used for training and the conditioning information. The best model for each environment is tabulated in Table A.1. The excess returns (alphas) are the estimate of the intercept of the SDF return series on the corresponding risk factors. The numbers in parentheses represent the t-statistic of the estimates. **Bold** means significance at the 5% level. All alphas are expressed in decimal points and represent monthly excess returns.

Predictive performance Table 5 shows the results of the beta net performance, as measured by the explained variation and the cross-sectional R^2.

Environment		**Results Test Set**	
Objective	Data	EV	XS-R^2
MP	Bondpf + Macro	0.64	0.88
SR	Bondpf + Macro	0.70	0.92
MP	Bondpf	0.40	0.71
SR	Bondpf	0.65	0.88

Table 5: Bond Portfolios - Predictive Performance

Notes: The environment describes the objective function that was used to train the mode and the conditioning information that the model had available. The best model for beta estimation for each environment is tabulated in Table A.2 in the appendix. The results show explained variation and cross-sectional R-squared as described in more detail in Section 3.4. All units are expressed in decimal points.

We see that Sharpe ratio maximization delivers somewhat higher predictability for the full feature set (portfolio characteristics and macro time series), and that excluding macro time series as predictive features lowers predictability for both SDF estimation approaches.

4.2 Individual bond results

The individual bonds contain 10,715 unique securities of close to 2,000 unique issuers. In principle, the granularity of the dataset allows for a more targeted portfolio allocation. While in theory this should have advantages, most of the literature in asset pricing has focused on the return characteristics of sorted portfolios due to the idiosyncratic variance at the individual security level. Advances in machine learning allow us to attempt the exercise at the security level.

Just as for bond portfolios, we estimate the SDF using both approaches: mispricing loss minimization and Sharpe ratio maximization. Besides the data environment with a full feature set, we also consider the environment where macro time series have been excluded. As an additional robustness exercise, we take the best architecture and hyperparameters estimated with a full feature set, and re-estimate it excluding bonds issued by financials as well as REITs. If the results do not change significantly based on this exclusion, we can assert that they are not driven by financials and REITs. Table 6 shows the SDF portfolio performance for each data environment and objective. The corresponding excess returns are shown in Table 7 in the text.

Environment		**Results Test Set**						
Objective	Data	Sharpe	Sortino	Calmar	Max-DD	CAGR	Max-lev	Mean-lev
MP	Ind. + Bond + Macro	0.17	6.30	0.02	0.20	0.04	3.15	1.29
SR	Ind. + Bond + Macro	0.27	3.93	0.03	0.27	0.12	4.62	2.42
MP	Ind. + Bond	0.12	0.46	0.02	0.56	0.08	4.66	3.51
SR	Ind. + Bond	0.29	2.54	0.05	0.29	0.16	4.29	2.58
MP	Ind. + Bond + Macro - Fin	0.25	2.83	0.03	0.28	0.10	2.70	2.19
SR	Ind. + Bond + Macro - Fin	0.26	2.02	0.03	0.39	0.14	3.97	3.20

Table 6: Individual Bonds - Performance

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. − Fin means financials and REITs were excluded. The best model for each environment is tabulated in Table A.3 in the appendix. The results show a set of financial ratios as described in more detail in Section 3.4, as well as maximum drawdown (Max-DD), and the portfolio implied maximum leverage (Max-lev) and mean leverage (Mean-lev). Sharpe, Sortino and Calmar ratios are monthly. All units are expressed in decimal points.

The best performing model with individual bonds achieves a monthly Sharpe ratio of 0.29, or annual Sharpe ratio of 1.00. Thus, the implied SDF portfolio shows even better return / risk characteristics relative to the best performing model for bond portfolios.

We find that maximizing the Sharpe ratio (Approach 2) performs overall better than minimizing the mispricing loss (Approach 1) in terms of Sharpe and Calmar ratios in the test set. Moreover, Sharpe ratio maximization leads to higher CAGR, together with higher and more significant excess returns.

Moreover, training models with Sharpe ratio maximization objective is always faster than with the mispricing loss. This discrepancy is considerably larger for individual bonds data than for bond portfolios. Hence, the computational bottleneck in the case of minimizing mispricing loss is most likely the calculation of the batch-mispricing loss (equation 5).

Figure 3 makes clear the dominance of the Sharpe ratio maximization approach in terms of cumulative returns over time. It depicts the value indices of the SDF portfolio in the test set for both objective functions, for the data environment without any exclusions of features or assets.

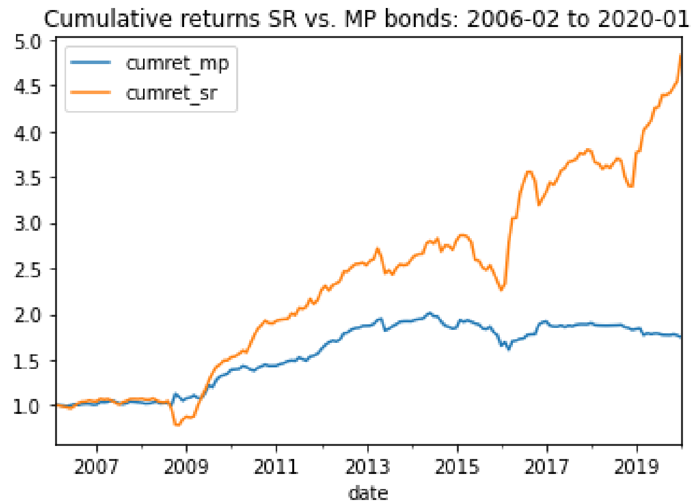


Figure 3: Ind. Bonds - Cumulative Return, full info set

The model with the highest Sharpe ratio in our estimation omits the macroeconomic time series as conditioning information. While this model achieves the highest Sharpe and Calmar ratios, the Sortino ratio is somewhat lower and drawdowns higher, together with higher mean leverage than the model estimated with the same objective but in an environment with macro features. We conjecture that the result of an improved Sharpe ratio under exclusion of macro features may not hold when searching across a larger hyperparameter search, e.g. for risk aversion parameters γ higher than 0.01 in the formulation of the linearized Sharpe ratio (see equation 7).

Finally, the results do not change qualitatively when excluding financials and REITs, though the discrepancy in terms of Sharpe and Sortino ratios becomes smaller.

Excess returns We test for excess returns against the same benchmarks as for bond portfolios. The model and objective combination with the highest Sharpe ratio yields monthly excess returns of 1.0% vis-a-vis the S&P 500, or 12.0% annually, with a t-statistic of 3.040. Even against the FF5 benchmark, the monthly excess return remains close to 1.0% with a t-statistic of 2.500 as shown in Table 7.

In principle, one could also consider the 40 bond portfolios as the risk factors that span the SDF. Without thorough study or precedent in the literature it is unclear which of the 40 bond portfolio represent risk factors, that is factors with non negative risk prices. We use all 40 portfolios as potentially relevant and find no statistically significant excess returns of the SDF portfolio with respect to these bond portfolios (see last column of Table 7). This suggests that the 40 bond portfolios span the SDF well.

Predictive performance Table 8 shows the predictive performance of the beta networks for individual bonds data. While the SDF portfolio estimated using the objective of Sharpe ratio maximization performs typically better in terms of financial ratios, the performance ordering reverses for predictability. For a given data environment, minimizing the mispricing loss leads to often substantial increases in the explained variation and the cross-sectional R^2. We did not observe this difference in predictive performance for bond portfolios – the predictive performance was better, while the Sharpe ratios were close to each other when

	Environment	Alphas in Test Set			
Objective	Data	S&P 500	FF 3 Factors	FF 5 Factors	Bond Portfolios
MP	Ind. + Bond + Macro	0.003	0.003	0.003	0.001
		(1.870)	(1.920)	(1.820)	(0.380)
SR	Ind. + Bond + Macro	**0.007**	**0.006**	**0.006**	0.001
		(2.700)	(2.380)	(2.290)	(1.000)
MP	Ind. + Bond	0.001	0.002	0.002	-0.001
		(0.180)	(0.530)	(0.480)	(0.710)
SR	Ind. + Bond	**0.010**	**0.008**	**0.009**	**0.003**
		(3.040)	(2.780)	(2.500)	(3.280)
MP	Ind. + Bond + Macro - Fin	**0.005**	**0.005**	**0.005**	0.000
		(2.410)	(2.100)	(1.930)	(0.700)
SR	Ind. + Bond + Macro - Fin	**0.008**	**0.007**	**0.007**	0.001
		(2.470)	(2.140)	(1.960)	(0.700)

Table 7: Individual Bonds - Excess Returns

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. − Fin means financials and REITs were excluded. The best model for each environment is tabulated in Table A.3 in the Appendix. The excess returns (alphas) are the estimate of the intercept of the SDF return series on the corresponding risk factors. The numbers in parentheses represent the t-statistic of the estimates. **Bold** means significance at the 5% level. All alphas are expressed in decimal points and represent monthly excess returns.

using the Sharpe ratio as the objective function vs. minimizing the mispricing loss. This difference may be a reflection on the differing idiosyncratic variation in the two datasets. The mispricing loss seems overall better suited to fit the model to the variation of the data when there is considerable idiosyncratic variation, whereas the Sharpe ratio maximization training objective allows for better generalization in terms of financial ratios in the test set.

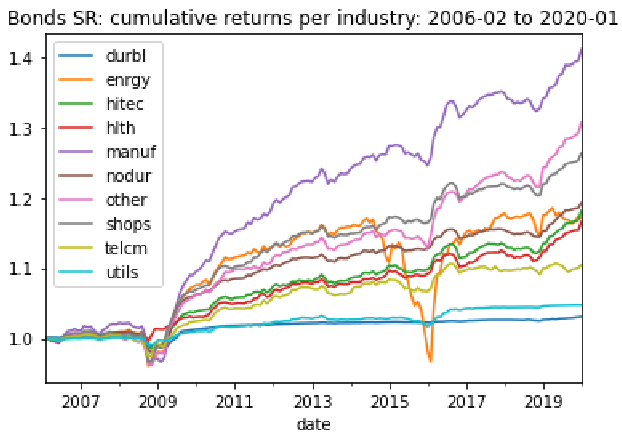
	Environment	Results Test Set	
Objective	Data	EV	XS-R^2
MP	Ind. + Bond + Macro	0.23	0.45
SR	Ind. + Bond + Macro	0.19	0.39
MP	Ind. + Bond	0.20	0.27
SR	Ind. + Bond	0.12	0.18
MP	Ind. + Bond + Macro - Fin	0.22	0.36
SR	Ind. + Bond + Macro - Fin	0.16	0.30

Table 8: Individual Bonds - Predictive Performance

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. − Fin means financials and REITs were excluded. The best model for each environment is tabulated in Table A.4 in the Appendix. The results show explained variation and cross-sectional R-squared as described in more detail in Section 3.4. All units are expressed in decimal points.

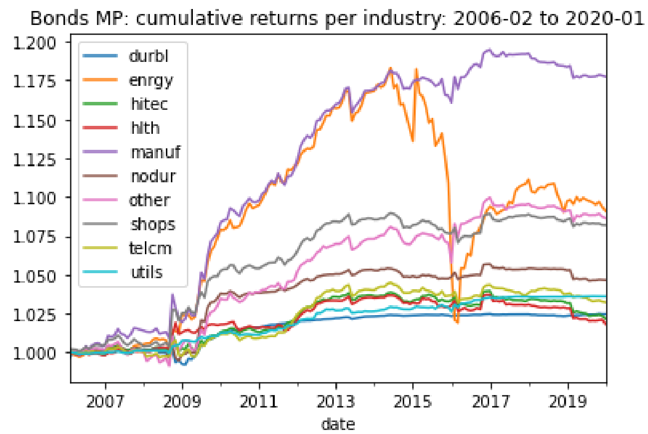
To further illustrate the difference between the two approaches, SR-max and MP-min, we calculate the split of the SDF portfolio return across industries in the test set. Figure 4 shows the cumulative returns in the test set for each separate industry.

The most pronounced difference is visible in the evolution of the energy sector (labeled "enrgy"). The investment into the energy sector sustains considerable losses in 2015 and early 2016 for both training objectives. This coincides with a drop in the oil prices from above \$100/barrel to about \$140/barrel. Overall, this had a larger influence on the returns under the mispricing loss objective, with persistent effects until the end of our sample period. This is because a greater degree of the cumulative return had come from the energy sector in the preceding years under this objective criterion. On the other hand, the cumulative return of the investment into the energy sector using the Sharpe ratio maximization objective only experienced a temporary decline with a subsequent bounce to the previous level. The overall contribution to the portfolio



(a) Industry Return - Sharpe Ratio

Notes: Cumulative returns per industry sector on test set for maximization of Sharpe ratio approach.



(b) Industry Return - Mispricing Loss

Notes: Cumulative returns per industry sector on test set for minimization of mispricing loss approach.

Figure 4: Cumulative Industry Returns

return is determined by the cumulative return of the industry times the exposure to the sector. In order to obtain the exposure to the energy sector, we calculate the signed leverage share of the energy sector for both models. Concretely, we calculate for each date the share of the total leverage due to the energy sector, and multiply it with the sign of the sum of weights in the respective sector:

$$sign(sum(\text{weights of energy})) \times \frac{sum(|\text{weights of energy}|)}{\text{total leverage}}.$$

The results of the calculation for the energy sector are shown in Figure 5.

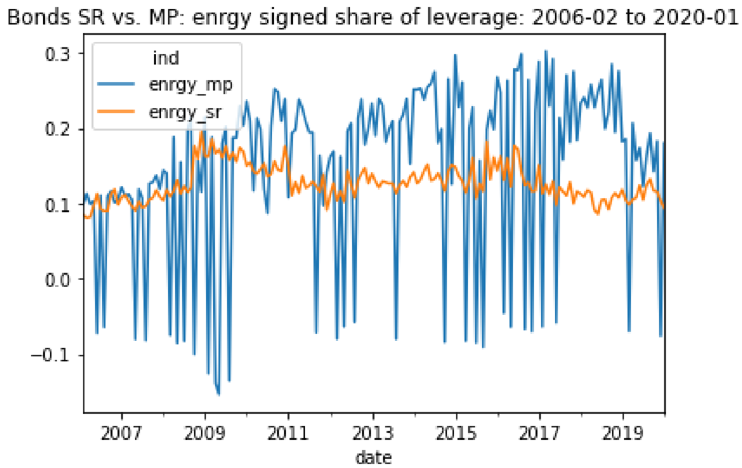


Figure 5: Energy sector: Signed share of leverage

The signed leverage share of the energy is much smoother when the objective is Sharpe ratio maximization. It also never crosses in negative territory and remains relatively stable over time. For the model that is estimated with the mispricing loss, the signed leverage shows large swings for the leverage share. For this training objective, we observe a similar pattern for the share of the overall portfolio leverage for several industries, but not all of them. In contrast, the signed leverage measure is always smoother for all industries in the case of Sharpe ratio maximization.

Ex-post analysis: Market regimes for the US corporate bond markets To gain additional insights into our model performance results, we estimate a Gaussian hidden Markov model (HMM) with two and three states on the 40 bond portfolios. This allows us to better understand the market regimes in the US corporate bond market, in a way that is decoupled from the results of our neural networks exercise. The Gaussian HMM simultaneously estimates the hidden states and the probability distribution across these states.

The HMM model with three states delivers states with the clearest interpretation.[9] Concretely, we find the following three states with the corresponding colors in parentheses:

A. **State 1:** high volatility, high positive returns (blue)

B. **State 2:** low volatility, low positive returns (orange)

C. **State 3:** high volatility, negative returns (green)

In Figure 6 we show the evolution of the equally-weighted value index on the 40 bond portfolios split according to the estimated HMM states.

The 1980s, a period with high inflation and large uncertainty about monetary policy, are characterized with a large presence of the high-volatility-negative return state. Starting in the early 1990s the low-volatility-small-positive-returns state begins to dominate. This dominance continues throughout the remainder of the sample period, albeit interspersed with clustered occurrences of the other two states. The market oscillates between the more extreme states 1 and 3 from the late 1990s to 2005. Throughout our test set, the dominant state is the low-volatility-low positive returns state, as Figure 6 shows. Interestingly, even the financial

[9] Section A.2 in the Appendix contains the numerical description of the hidden states.

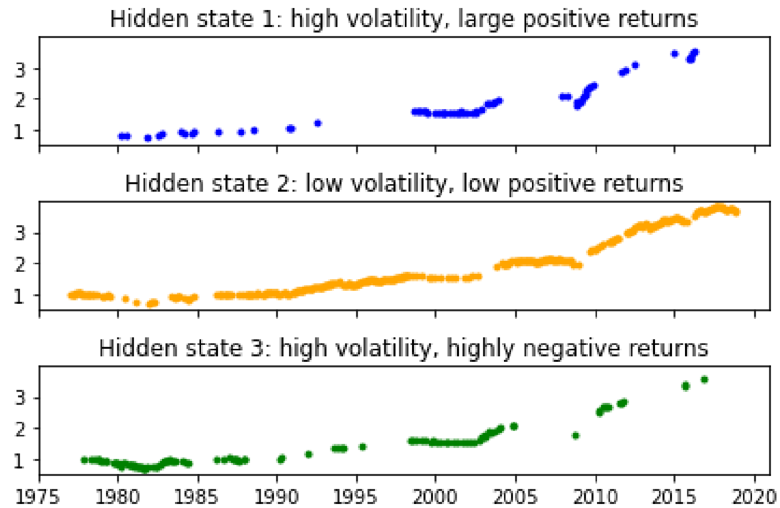


Figure 6: Bond Portfolios - HMM state evolution vs. equally-weighted bond portfolio index.

dot.com bust and the financial crisis of 2008/2009 does little to change the dominance of the low-volatility-low-positive returns state. At first, this might be surprising but it is important to keep in mind that the bonds in our sample are the most liquid bonds, a predominant share of which are issued by companies in the S&P500 as described in Section 2.4.

5 Contrast to literature and other models

In 5.1 we present a more "traditional" asset pricing benchmark in the form of a linear pricing kernel and dimensionality reduction of the candidate macro predictors via a dynamic factor model, which we estimate via general method of moments (GMM). Section 5.2 presents an alternative approach to estimation that is based on the modeling in Chen et al. (2023). Its focus is again minimizing deviations from the conditional no arbitrage requirement in equation (1), but now it uses conditional test assets for the estimation.

5.1 Linear Models

Using characteristic sorted portfolios as test assets has a long history in the empirical asset pricing literature for equity. It comes at the advantage of having a consistent time series for each portfolio, when the latter contains assets that are substituted dynamically over time. Further it reduces some of the idiosyncratic variance of individual assets. However, the formation of portfolios prevents us from using individual firm or bond characteristics as candidate predictors. We thus focus on macroeconomic aggregate time series to construct candidate test assets.

Dynamic factor model. The large number of potential candidate predictors that are observable in the macroeconomy as well as the limited amount of test-assets, requires a dimensionality reduction. One popular way to do this is to estimate a dynamic factor model and to use the factors as candidate predictors. We use a dynamic factor model with the following structure:

$$x_{it} = \lambda_i' F_t + \epsilon_{it}, \qquad i = 1, \ldots, N', \quad t = 1, \ldots, T'.$$

where, F_t is a $r \times 1$ vector (the latent factors) and λ_i are the latent factor loadings. One of the choices that a dynamic factor model requires is the optimal number of factors r. We use the information criteria as presented in Bai and Ng (2002) to determine the optimal number of factors in our setting. Specifically, given the large number of macro time series when compared to the number of periods T, we use their information criteria $PC_{pi}, i = 1, 2, 3$ and IC_{p3}. We evaluate the robustness of the optimal number of factors by evaluating versions of these information criteria where we introduce a hyperparameter that multiplies the penalty for model complexity, and we apply these criteria where the penalty for model complexity/overfitting is scaled down by respectively $10, 20, 30\%$.[10] We find that 45 macro factors is the most sensible suggestion according to the information criteria of Bai and Ng (2002). While 45 factors is a substantial reduction from the original 132 factors, it is still large relative to the 40 available test portfolios. We follow Ludvigson and Ng (2009) by imposing an ex-ante restriction on the number of factor; in our case a maximum of 39 factors. With this restriction, the $PC_{pi}, i = 1, 2, 3$ and IC_{p3} criteria give an optimal number of 30 factors. We also test more parsimonious models in accordance with Ludvigson and Ng (2009). In particular, we consider specifications with 7, 8, and 9 factors.

Asset Pricing. We first investigate whether macroeconomic information helps in pricing U.S. corporate bonds. We postulate that the stochastic discount factor (SDF) depends linearly on a number of macro factors and the Cochrane-Piazessi (CP) factor as first defined in Cochrane and Piazessi (2008).[11]

$$sdf_t = 1 - \beta' F_t - \gamma' CP_t, \tag{10}$$

where $F_t = [F_{1t}, \ldots, F_{(m-1)t}]$, with $m = 8, 9, 10, 30$. Here, m is the number of factors of the dynamic factor model and CP_t is the Cochrane-Piazessi factor.

We estimate the stochastic discount factor via general methods of moments (GMM) where the asset pricing restrictions for each test asset provide the moment restrictions. Thus, we have 40 moment restrictions to estimate the SDF. Given that the set of moment restrictions is strictly larger than the number of parameters to estimate, we can perform over-identification tests (J-test) for the statistical validity of the SDF.

Table 9 shows the results of the estimation. In particular we test all combinations of $m = 8, 9, 10, 30$ factors and the inclusion or omission of the Cochrane-Piazessi factor.

In terms of explainability of the variance, the addition of macro factors to the CP factor increases R^2 and the explained variance (EV) by orders of magnitude. This is qualitatively true regardless of the number of the macro factors included. The addition of the CP factor to specifications with macro factors only, improves R^2 and EV, albeit sometimes by a small amount. The CP factor is highly significant in all specifications. The R^2 value of the CP-only regression is much lower than in the exercise of Ludvigson and Ng (2009). This suggests that a lot of the variation for corporate bonds is not due to the CP factor, even though the factor is highly significant. Otherwise, with macro PCA factors, R^2 are not orders of magnitude different from Ludvigson and Ng (2009).

The J-test for the violation of the moment conditions also has small p-value with CP only, but p-value increases significantly once macro PCA factors are considered. For every specification with macro factors and the CP factor, a J-test testing the null-hypothesis that the difference between the unrestricted model containing macro factors and the restricted model with only the CP factor is small, delivers a p-value that is

[10] The scaling of the penalty function for model complexity does not change the formal results of Bai and Ng (2002). In particular, the versions of PC/IC that we consider for robustness, are all valid information criteria for the selection of the number factors within the framework of Bai and Ng (2002).

[11] See also Hodrick and Tomunen (2018) for the estimation of the CP factor.

Environment		Results				
Macro PCA	CP	Sharpe	EV	R^2	J-test p-val all moments	J-test p-val macro+CP vs. CP
0	1	-0.09	0.04	0.02	0.01	n/a
7	0	0.10	0.37	0.35	0.57	n/a
	1	0.07	0.37	0.36	0.60	0.00
8	0	0.07	0.37	0.36	0.53	n/a
	1	0.05	0.37	0.36	0.56	0.00
9	0	0.08	0.38	0.36	0.47	n/a
	1	0.06	0.38	0.37	0.69	0.00
30	0	0.04	0.54	0.53	0.97	n/a
	1	-0.03	0.54	0.53	0.99	0.00

Table 9: Performance Measures GMM

Notes: The environment describes the number of macro PCA factors used in the GMM specification (10), and whether the Cochrane-Piazessi factor was included. The results show the Sharpe ratio, the explained variation, and the p-values for a J-test of all moments and of macro PCA + CP vs. CP. For the definition of explained variation see Section 3.4. All units are expressed in decimal points.

indistinguishable from 0 to more than ten decimal places. This is evidence that the specification containing the CP factor only is an insufficient model for the bond portfolios, whereas including multiple macro PCA factors increases the explainability of the model.

We also see that, even though the estimation is in-sample, the Sharpe ratios and explained variations are significantly lower than the ones achieved by our neural network estimation (compare with Tables 3 and 5).

Predictive performance. We further test the predictive content of macroeconomic time series in predictive regression. In particular, we estimate specifications of the form:

$$rx_{j,t+1} = \alpha_j + \beta_j' F_t + \gamma_j CP_t + \epsilon_{j,t+1}, \tag{11}$$

where $rx_{j,t+1}$ is the one month excess log return on the bond portfolio j, F_t is a vector of factors that are estimated via principal component analysis and CP_t is the Cochrane-Piazessi factor as constructed in Cochrane and Piazessi (2008). We consider here specifications with 8 macro PCA factors.

Predictive regressions in-sample. Tables A.14-A.16 in the appendix present the result of in-sample predictive regressions for our 40 bond portfolios. In comparison to Cochrane and Piazzesi (2005) the CP factor has lower predictive power than in the treasury bond market. The R^2 on specifications that contain only the CP factor averages about 0.03. We find that macro factors improve the explanatory power of risk premiums substantially as evidenced by R^2-s of about 0.10 with 8 factors and the CP factor. This is still substantially lower than the findings in Ludvigson and Ng (2009) for treasury bonds with 5 PCA factors and the CP factor in the sample period 1964:1-2003:12.[12]

Predictive regressions out-of-sample Tables A.17-A.18 in the appendix report the results for predictive regression that are strictly out of sample. The R^2 on specifications that contain only the CP factor averages

[12] Ludvigson and Ng (2009) find R^2s of 0.40 and above for the 5 factors + CP factor specifications. However, using the 5 factors alone brings the R^2 down to between 0.14 and 0.26. The discrepancy to our results can be mainly attributed to the differential explanatory power of the CP factor. This is not too surprising since the CP factor has been estimated in the treasury bond market. In an extension one could think about constructing an analogous factor for the corporate bond market.

to slightly less than 0.03 with typical values that are slightly lower than the in-sample case. We again find that macro factors improve the explanatory power of risk premiums substantially as evidenced by R^2-s of about 0.09 with 8 factors and the CP factor.

Finally, we perform a Diebold-Mariano test to compare residuals resulting from the beta estimation via neural networks as discussed in section 4.1 with residuals from the predictive regressions. Table 10 shows the Diebold-Mariano statistic for the comparison between residuals from the Sharpe ratio maximization approach vs. predictive regressions and residuals from the from the mispricing loss minimization vs. predictive regressions. Both test statistics are significant at the 1% level, and we see the dominance of the nonlinear methods via neural networks in general.[13]

Comparison	DM statistic
SR vs. predReg	**-4.82**
MP vs. predReg	**-4.82**

Table 10: Bond Portfolios - Predictive performance of predictive regressions vs. SR, MP.

Notes: The table depicts the results of a Diebold-Mariano test of the residuals from the predictive regressions based on the specification (10) with the Cochrane-Piazessi factor included, versus those coming from the beta-estimation with objectives respectively Sharpe-ratio maximization, and mispricing loss minimization. Both results are significant at the 1% level. All units are expressed in decimal points.

5.2 Alternative neural network architectures

A generalization of the mispricing loss (5) consists of using the full conditional expectation in the no-arbitrage condition of equation (1). One considers test assets that depend on the information available at time t, and the respective mispricing loss as a training criterion. Chen et al. (2023) uses a GAN-like (generative adversarial network) architecture to estimate an asset pricing model of the U.S. equity market based on this generalized moment condition. They focus on the following optimization problem.

$$\min_{\omega(\cdot)} \max_{g(\cdot)} \frac{1}{N} \sum_{i=1}^{N} \frac{|T_i|}{T} \left\| \frac{1}{|T_i|} \sum_{t \in |T_i|} \left(1 - \omega(z_{t,\cdot}) \cdot R^e_{t+1}\right) R^e_{t+1,i} g(z_{t,i}) \right\|^2 . \tag{12}$$

Here, $g(z_{i,t})$ are potentially tradeable test assets that depend on features $z_{i,t}$ known at time t for asset i. They act like a critic to the construction of weights $\omega(z_{t,\cdot})$. The mispricing loss (5) we use can be seen as a simplification of (12), where the inner optimization does not appear and the test assets are chosen to be constant with respect to features $z_{i,t}$. Chen et al. (2023) find better performance of the SDF portfolio on U.S. equities, relative to other model candidates that their paper considers. Their GAN architecture consists of two networks, one producing the SDF weights ω, and one producing the test assets g. The networks are jointly trained in an end-to-end fashion.

We have implemented a version of their approach within our set up of Pytorch, with a few differences in terms of the coding approach that are detailed in the appendix, section A.3. The main difference in approach is that we add to the features a double index to identify the asset over time, as explained in section 2.4 on data management. This allows using the batch size as a hyperparameter for the Adam optimizer, allows for less memory usage, and keeps the information used by the algorithm to compute the weights at a minimum,

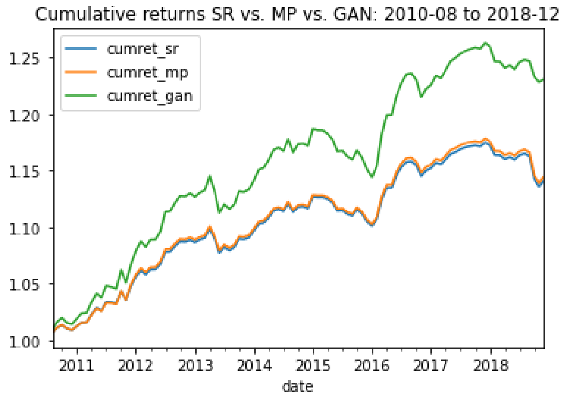
[13] This comparison uses residuals for the whole sample. The numbers are qualitatively similar if one focuses instead on the time period of the test set of the neural networks exercise.

thus avoiding look-ahead bias. Because the training of the GAN architecture takes considerably more time than the architectures that are the main focus of our paper, we have done only a limited search over the hyperparameter space of the optimization and over the possible architecture parameters.

We have tested our version of the GAN architecture for both individual bonds and bond portfolios. We find that for individual bonds a form of *mode collapse* appears: the SDF portfolio weights ω are equal, or close-to-equal across assets for every time period. Hence the predicted SDF is a multiple of the equally weighted portfolio. We claim that this issue is not inherent to the architecture, and should disappear with a greater hyperparameter/architecture search.

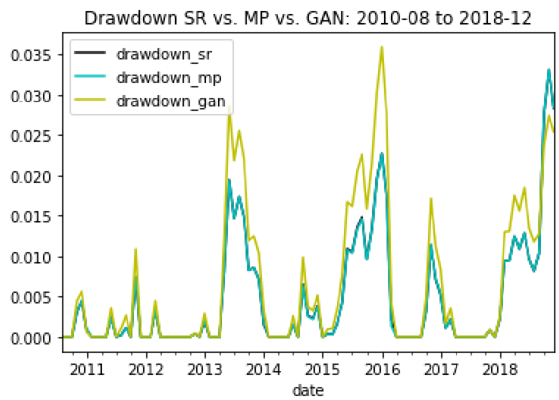
For bond portfolios, the best model from the hyperparameter and architecture search does not exhibit the mode collapse phenomenon. It achieves a high monthly Sharpe ratio of .30, as well as higher Calmar and Sortino ratios than the Sharpe ratio maximization and mispricing loss minimization approach on the intersection of the test sets.[14] In terms of cumulative returns, it outperforms the other models as Figure 7a shows. Together with the higher Sharpe-ratios and higher cumulative returns, the model shows higher mean leverage and somewhat higher drawdowns (see Figure 7b) than the Sharpe ratio maximization approach. The excess returns with respect to the S&P500, FF3, and FF5 are of the same magnitude as for the approach with the Sharpe ratio maximization, but with slightly lower significance levels.

[14]Note that the test sets are different, because the GAN network uses RNNs which take as input sequences of macro data. Hence the test set in the GAN case is smaller.



(a) Cumulative returns of the three approaches.

Notes:sr stands for Sharpe ratio maximization, *mp* stands for mispricing loss minimization, *gan* stands for GAN.



(b) Drawdown of the three approaches.

Notes: sr stands for Sharpe ratio maximization, *mp* stands for mispricing loss minimization, *gan* stands for GAN.

Figure 7: GAN cumulative returns and drawdown of the two main approaches vs. GAN.

As a separate test, we have compared the GAN approach vs. the Sharpe ratio maximization approach with the 10 industry portfolios from Kenneth French's website (monthly returns). We find again that, compared to Sharpe ratio maximization, GAN delivers higher Sharpe ratio and overall higher CAGR over time, but this time at a higher maximum drawdown and thus lower Calmar and Sortino ratios.

More details on the comparisons can be found in the Appendix Section A.3.

6 Conclusion

We build on recent advances in data availability and machine learning to estimate asset pricing models and document their out of sample performance for the U.S. corporate bonds market. We deploy neural network architectures to flexibly model the stochastic discount factor for two types of datasets. One dataset consists of bond portfolios, analogous to most of the literature for asset pricing of equities. The second dataset consists of individual bond level data, that we augment with financial fundamentals of the bond issuer. In the latter case, we estimate asset pricing models for corporate bonds at the individual security level, similar to recent studies that estimate asset pricing models at the security level for equities.

As the guiding principle for our estimation, we use two approaches that, under the Efficient Market

Hypothesis, are implied by no arbitrage conditions. The first approach minimizes the empirical pricing errors as the objective function in the estimation. The second approach maximizes the Sharpe ratio of the stochastic discount factor to estimate the pricing model. Despite the equivalence of the two approaches in theory, for the case of individual bonds, we find diverging performance of the resulting asset pricing model in our setting. The approach that maximizes the Sharpe ratio delivers better out-of-sample performance in terms of Sharpe and Calmar ratios and often larger and statistically stronger excess returns relative to conventional asset pricing benchmarks/risk factors. For individual bonds data, we find that using the mispricing loss leads generally to higher predictive performance. The discrepancies between the two approaches are smaller for portfolio level data. We contrast these non-linear approaches to the estimation of the SDF with conventional approaches that postulate linearity of the SDF in its pricing factors. We find that the non-linear methods lead to smaller pricing errors, higher Sharpe ratios and more variance explainability.

Our estimation approach using mispricing loss does not consider test assets. While we consider an alternative based on the work in Chen et al. (2023) on estimation of the SDF for equities (GAN network), that takes into account endogenously produced test assets, we have not experimented enough with it to replicate its success with equities data in our setup with individual bonds data. We show that for bond portfolios, the GAN approach leads to higher Sharpe ratios, whereas for individual bonds a form of mode collapse appears: the GAN-implied SDF is a multiple of the equally-weighted portfolio. Moreover, estimating the GAN architecture comes at considerably higher computation costs. More work is needed to compare the different architectures.

Furthermore, given that Sharpe ratio maximization performs better than the mispricing loss in our set up of non-conditional test assets, it would be interesting to pursue whether an intermediate approach is feasible, where Sharpe ratio maximization is the goal, but there is a critic similar to the GAN network, that tests the estimated SDF using test assets constructed from data. The paper Bryzgalova et al. (2020) goes in that direction in the case of equities using tree algorithms instead of neural networks.

Finally, another avenue for future research would be to estimate a joint asset pricing model for both the U.S. corporate bonds and the U.S. equity market. Our work creates a link between bond issuer equity characteristics and their issued corporate bonds. Jointly estimating the SDF on both asset classes would contribute to a better understanding of the interplay of the corporate bonds and equity markets.

References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Back, K. E. (2017). *Asset Pricing and Portfolio Choice Theory, 2nd ed.* New York: Oxford University Press.

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

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Balakrishnan, K., Bartov, E., and Faurel, L. (2010). Post loss/profit announcement drift. *Journal of Accounting and Economics*, 50(1):20–41.

Bali, T. G., Subrahmanyam, A., and Wen, Q. (2021). Long-term reversals in the corporate bond market. *Journal of Financial Economics*, 139(2):656–677.

Ball, R., Gerakos, J., Linnainmaa, J. T., and Nikolaev, V. V. (2015). Deflating profitability. *Journal of Financial Economics*, 117(2):225–248.

Basu, S. (1983). The relationship between earnings' yield, market value and return for nyse common stocks: Further evidence. *Journal of financial economics*, 12(1):129–156.

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The journal of finance*, 43(2):507–528.

Bianchi, D., Büchner, M., and Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies*, 34(2):1046–1089.

Bryzgalova, S., Pelger, M., and Zhu, J. (2020). Forest through the trees: Building cross-sections of stock returns. *Available at SSRN 3493458*.

Bustamante, M. C. and Donangelo, A. (2017). Product market competition and industry returns. *The Review of Financial Studies*, 30(12):4216–4266.

Chen, L., Pelger, M., and Zhu, J. (2023). Deep learning in asset pricing. *Management Science*.

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

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

Chung, K. H. and Zhang, H. (2014). A simple approximation of intraday spreads using daily data. *Journal of Financial Markets*, 17:94–120.

Cochrane, J. (2009). *Asset pricing: Revised edition*. Princeton university press.

Cochrane, J. H. and Piazessi, M. (2008). Decomposing the yield curve. *University of Chicago working paper, January 2008 version.*

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

Cooper, M. J., Gulen, H., and Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *the Journal of Finance*, 63(4):1609–1651.

Culp, C. L., Nozawa, Y., and Veronesi, P. (2018). Option-based credit spreads. *American Economic Review*, 108(2):454–88.

Daniel, K., Mota, L., Rottke, S., and Santos, T. (2020). The cross-section of risk and returns. *The Review of Financial Studies*, 33(5):1927–1979.

Datar, V. T., Naik, N. Y., and Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of financial markets*, 1(2):203–219.

De Bondt, W. F. and Thaler, R. (1985). Does the stock market overreact? *The Journal of finance*, 40(3):793–805.

Desai, H., Rajgopal, S., and Venkatachalam, M. (2004). Value-glamour and accruals mispricing: One anomaly or two? *The Accounting Review*, 79(2):355–385.

D'Acunto, F., Liu, R., Pflueger, C., and Weber, M. (2018). Flexible prices and leverage. *Journal of Financial Economics*, 129(1):46–68.

Elkamhi, R., Jo, C., and Nozawa, Y. (2020). A one-factor model of corporate bond premia. *Available at SSRN.*

Elkamhi, R., Jo, C., and Nozawa, Y. (2021). A one-factor model of corporate bond premia1. *Working paper, January 2021 version.*

Fama, E. and Bliss, R. R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77:680–692.

Fama, E. and French, K. (1992). The cross-section of expected stock returns. *Journal of Finance*, 47(2):427–65.

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

Fama, E. and French, R. K. (2015a). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

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

Fama, E. F. and French, K. R. (1995). Size and book-to-market factors in earnings and returns. *The journal of finance*, 50(1):131–155.

Fama, E. F. and French, K. R. (2012). Size, value, and momentum in international stock returns. *Journal of financial economics*, 105(3):457–472.

Fama, E. F. and French, K. R. (2015b). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Fama, E. F., French, K. R., et al. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance*, 51(1):55–84.

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

Freyberger, J., Neuhierl, A., and Weber, M. (2020a). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377.

Freyberger, J., Neuhierl, A., and Weber, M. (2020b). Dissecting characteristics nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377.

Gandhi, P. and Lustig, H. (2015). Size anomalies in us bank stock returns. *The Journal of Finance*, 70(2):733–768.

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

George, T. J. and Hwang, C.-Y. (2004). The 52-week high and momentum investing. *The Journal of Finance*, 59(5):2145–2176.

Gorodnichenko, Y. and Weber, M. (2016). Are sticky prices costly? evidence from the stock market. *American Economic Review*, 106(1):165–99.

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

Gu, S., Kelly, B., and Xiu, D. (2021). Autoencoder asset pricing models. *Journal of Econometrics*, 222(1):429–450.

Guijarro-Ordonez, J., Pelger, M., and Zanotti, G. (2021). Deep learning statistical arbitrage. *arXiv preprint arXiv:2106.04028*.

Haugen, R. A. and Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of financial economics*, 41(3):401–439.

He, X., Feng, G., Wang, J., and Wu, C. (2021). Predicting individual corporate bond returns. *Available at SSRN 4374213*.

He, Z., Kelly, B., and Manela, A. (2017). Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35.

Hendrycks, D. and Gimpel, K. (2016). Gaussian error linear units (gelus). *arXiv preprint arXiv:1606.08415*.

Hirshleifer, D., Hou, K., Teoh, S. H., and Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, 38:297–331.

Hodrick, R. and Tomunen, T. (2018). Taking the cochrane-piazzesi term structure model out of sample: More data, additional currencies, and fx implications. *NBER Working paper, September 2018 version.*

Hou, K., Karolyi, G. A., and Kho, B.-C. (2011). What factors drive global stock returns? *The Review of Financial Studies*, 24(8):2527–2574.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91.

Kaldor, N. (1966). Marginal productivity and the macro-economic theories of distribution: Comment on samuelson and modigliani. *The Review of Economic Studies*, 33(4):309–319.

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

Kingma, D. and Ba, J. (2015). Adam: A method for stochastic optimization. *ICLR 2015*.

Koijen, R. S. and Van Nieuwerburgh, S. (2011). Predictability of returns and cash flows. *Annu. Rev. Financ. Econ.*, 3(1):467–491.

Lewellen, J. (2015). The cross-section of expected stock returns. *Critical Finance Review*, 4(1):1–44.

Li, X., Chen, S., Hu, X., and Yang, J. (2019). Understanding the disharmony between dropout and batch normalization by variance shift. *CVPR 2019*, pages 2682–2690.

Litzenberger, R. H. and Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of financial economics*, 7(2):163–195.

Ludvigson, S. C. and Ng, S. (2009). Macro factors in bond risk premia. *The Review of Financial Studies*, 22(12):5027–5067.

Lyandres, E., Sun, L., and Zhang, L. (2008). The new issues puzzle: Testing the investment-based explanation. *The Review of Financial Studies*, 21(6):2825–2855.

McCracken, M. W. and Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2):449–470.

Novy-Marx, R. (2011). Operating leverage. *Review of Finance*, 15(1):103–134.

Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics*, 103(3):429–453.

Nozawa, Y. (2017). What drives the cross-section of credit spreads?: A variance decomposition approach. *Journal of Finance*, 72(5):2045–2072.

Palazzo, B. (2012). Cash holdings, risk, and expected returns. *Journal of Financial Economics*, 104(1):162–185.

Pontiff, J. and Woodgate, A. (2008). Share issuance and cross-sectional returns. *The Journal of Finance*, 63(2):921–945.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting review*, pages 289–315.

Soliman, M. T. (2008). The use of dupont analysis by market participants. *The Accounting Review*, 83(3):823–853.

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

Appendix

A.1 Best models for SDF estimation and beta estimation

A.1.1 Bond portfolios

Environment		Best model				
Objective	Data	Architecture	bs	lr	l2reg	l1reg
MP	Ind. + Bond + Macro	Model 1	256	0.005	0.001	none
SR	Ind. + Bond + Macro	Model 3	256	0.005	none	none
MP	Ind. + Bond	Model 2	256	0.001	0.001	none
SR	Ind. + Bond	Model 2	512	0.001	none	0.001

Table A.1: Bond Portfolios - Best Models for SDF estimation

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. The best model describes the architecture, batch size (bs), learning rate (lr) and the l1 (l1reg) and l2 (l2reg) regularization for the model that obtained the highest Sharpe ratio in the validation set.

Environment		Best model				
Objective	Data	Architecture	bs	lr	l2reg	l1reg
MP	Ind. + Bond + Macro	Model 2	256	0.0001	0.1	0.1
SR	Ind. + Bond + Macro	Model 3	256	0.0001	0.1	0.1
MP	Ind. + Bond	Model 4	256	0.0001	0.1	0.1
SR	Ind. + Bond	Model 3	256	0.0001	0.1	0.1

Table A.2: Bond Portfolios - Best Models for Beta estimation

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. The best model describes the architecture, batch size (bs), learning rate (lr) and the L^1 (l1reg) and L^2 (l2reg) regularization for the model that obtained the lowest mean-squared error the validation set.

A.1.2 Individual bonds

Environment			Best model				
Objective	Data	Exclusions	Architecture	bs	lr	l2reg	l1reg
MP	Ind. + Bond + Macro		Model 1	1024	0.0001	none	0.001
SR	Ind. + Bond + Macro		Model 3	512	0.0001	none	0.001
MP	Ind. + Bond		Model 4	512	0.005	0.001	none
SR	Ind. + Bond		Model 4	2048	0.005	0.001	0.001
MP	Ind. + Bond + Macro	No financials	Model 1	1024	0.0001	none	0.001
SR	Ind. + Bond + Macro	No financials	Model 3	512	0.0001	none	0.001

Table A.3: Individual Bonds - Best Model for SDF estimation

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. The best model describes the architecture, batch size (bs), learning rate (lr) and the L^1 (l1reg) and L^2 (l2reg) regularization for the model that obtained the highest Sharpe ratio in the validation set.

Environment			Best model				
Objective	Data	Exclusions	Architecture	bs	lr	l2reg	l1reg
MP	Ind. + Bond + Macro		Model 1	256	0.0001	none	0.1
SR	Ind. + Bond + Macro		Model 4	512	0.0001	0.1	0.001
MP	Ind. + Bond		Model 2	512	0.005	0.001	none
SR	Ind. + Bond		Model 2	1024	0.0001	0.1	0.1
MP	Ind. + Bond + Macro	No financials	Model 1	256	0.0001	none	0.1
SR	Ind. + Bond + Macro	No financials	Model 4	512	0.0001	0.1	0.001

Table A.4: Individual bonds - Best Models for Beta estimation

Notes: The environment describes the objective function that was used to train the model and the conditioning information that the model had available. The best model describes the architecture, batch size (bs), learning rate (lr) and the L^1 (l1reg) and L^2 (l2reg) regularization for the model that obtained the lowest mean-squared error the validation set.

A.2 Numerical results of the hidden-Markov model estimation

We used the package hmmlearn to perform the Gaussian hidden Markov model estimation. We estimate the volatility of the state by calculating the largest eigenvalue of the corresponding estimated covariance matrix. The estimated means and largest eigenvalue of estimated covariance calculated by the algorithm are as follows.

- **State 1:**

$$\text{means} = \begin{bmatrix} 0.81 & 1.29 & 1.61 & 1.90 & 2.52 & 1.23 & 1.34 & 1.52 & 1.53 & 2.00 \\ 0.80 & 1.26 & 1.80 & 1.81 & 2.32 & 1.00 & 1.43 & 1.73 & 1.77 & 2.29 \\ 1.61 & 1.57 & 1.67 & 1.66 & 1.93 & 1.28 & 1.27 & 1.49 & 1.32 & 2.93 \\ 1.19 & 2.28 & 1.46 & 1.59 & 1.53 & 1.59 & 1.65 & 1.64 & 1.56 & 1.55 \end{bmatrix}$$

$$\text{largest eigenvalue of covariance} = 277.05.$$

- **State 2:**

$$\begin{bmatrix} 0.18 & 0.24 & 0.24 & 0.21 & 0.31 & 0.22 & 0.21 & 0.23 & 0.25 & 0.33 \\ 0.17 & 0.23 & 0.24 & 0.23 & 0.34 & 0.19 & 0.25 & 0.24 & 0.27 & 0.22 \\ 0.18 & 0.22 & 0.22 & 0.29 & 0.31 & 0.24 & 0.21 & 0.24 & 0.22 & 0.25 \\ 0.17 & 0.44 & 0.17 & 0.19 & 0.18 & 0.19 & 0.26 & 0.26 & 0.32 & 0.34 \end{bmatrix}$$

$$\text{largest eigenvalue of covariance} = 46.64.$$

- **State 3:**

$$\begin{bmatrix} -0.20 & -0.41 & -0.54 & -0.55 & -0.36 & -0.23 & -0.32 & -0.43 & -0.49 & -0.56 \\ -0.14 & -0.40 & -0.52 & -0.51 & -0.45 & -0.12 & -0.36 & -0.53 & -0.62 & -0.57 \\ -0.50 & -0.49 & -0.51 & -0.54 & -0.21 & -0.37 & -0.38 & -0.40 & -0.43 & -0.21 \\ -0.53 & 0.44 & -0.45 & -0.43 & -0.38 & -0.30 & -0.42 & -0.41 & -0.34 & -0.31 \end{bmatrix}$$

$$\text{largest eigenvalue of covariance} = 262.82.$$

A.3 Alternative architectures: GAN

Chen et al. (2023) estimate the SDF portfolio F_{t+1} for the U.S. stock market using as features individual stock characteristics, as well as a large set of macro time series.[15] Their feature set is similar to the individual issuer characteristics and the macro data we use. The generalized mispricing loss function they use for the estimation is discussed in section 5.2 (see equation 12). If not familiar with the their paper, the reader should first refer to that section.

The weights for the test assets g_t are modeled as a neural network function of the same set of individual and aggregate characteristics as the weights ω_t (Moments Network). The weights of the SDF portfolio are produced from a separate neural network with the same structure (SDF network). In both networks, the macro data are passed through LSTM layers that extract a macro state, the individual characteristics through a FFN. After this step, the two outputs, individual characteristics as well as the macro state, are combined

[15] The code for their paper is available at https://github.com/jasonzy121/Deep_Learning_Asset_Pricing.

through another sequence of fully connected FFN layers to produce the final weights of the portfolio. The two networks, SDF and Moments network are jointly optimized in the following way:

- **Step 1:** run for a certain number of epochs the generalized mispricing loss but with test assets $g_t \equiv 1$. The training objective corresponds to the mispricing loss we use in the paper (5).
- **Step 2:** run for a certain number of epochs the maximization part of (12), i.e. produce the test assets g.
- **Step 3:** run for a certain number of epochs the minimization part of (12), i.e. produce the final portfolio weights.

This procedure is only repeated once – one global epoch. Moreover, because they do not split the train, validation and test datasets into mini-batches, each of the train-, validation-, test datasets is fed to the algorithm in one respective batch.

Note that the stocks dataset is also inherently imbalanced over time. Chen et al. (2023) deal with this unbalancedness by feeding a balanced dataset in one batch, where the missing returns are replaced by a dummy value UNK.

We have implemented a more online and more memory efficient version of their algorithm. Our version of the GAN algorithm keeps the same structure of the SDF and Moments networks, but has the following differences to the original version:

- We can work with the batch size as a hyper-parameter for the Adam optimizer.
- We feed the dataset without bloating it through replacing missing assets with a dummy UNK value.
- We allow for repetition of the trainings Steps 1-3 above, i.e. we allow training with a variable number of global epochs.
- We work with GRU cells as an RNN network for the macro time series instead of LSTM cells. GRU cells are typically faster at processing inputs than the LSTM network.
- We feed to the RNN of choice sequences of 12 months of macro data. According to a Ljung-Box test we performed on the partial auto-correlation function of the macro data, 12 months are more than the last significant lag for the macro time series.
- In our Pytorch setup, we need to set up two different optimizers, one for the SDF network, and one for the Moments network. We use the same parameters for both optimizers, including learning rate, i.e. we do not differentiate the hyperparameters of the optimizers based on the type of network.

The first two differences in our approach as listed above, are a consequence of the double index approach for data management that is explained in section 2.4.

The space of the possible architectures is much larger for the GAN approach than for the Sharpe ratio maximization and mispricing loss minimization approaches. Given computational constraints, we have tried two different architecture definitions with the GAN approach for SDF estimation. They are distinguished by the FFN subnetwork used in the SDF network to merge the macro hidden states with the individual characteristics of the assets. In both architectures the macro states, for each of the SDF and Moments networks, are produced via a GRU network with one hidden layer and a dimension of the hidden macro state of 6. The latter is the same choice as in Chen et al. (2023).

Specifically, we consider the following models for the SDF network.

- **Model 1 - SDF network:** individual characteristics and the macro state produced by the GRU network are passed together into one FFN network with 4 hidden layers and hidden units $[64, 32, 16, 8]$.

- **Model 2 - SDF network:** individual characteristics are processed first through two FFN layers with hidden units $[32, 16]$. Its output is combined with the macro state produced by the GRU network as an input for a FFN with 6 layers with hidden units $[64, 64, 32, 32, 16, 8]$.

For the Moments network we consider only one configuration given by

- **Moments network:** individual characteristics and the macro state produced by the GRU network are passed together as input for a FFN network with 3 hidden layers and hidden units $[32, 16, 8]$. The moments network outputs 6 assets, the same value for this hyperparameter as in Chen et al. (2023).

Regularization follows the same structure as the two approaches considered in the main paper. In particular, we regularize by using batch-normalization in the SDF and Moments networks, as well as L^1-regularization, L^2-regularization. Finally, we again use a dropout rate of 0.01.

We split the datasets that we use in the same way as for the approaches in the main paper: 42% train, 28% validation, and 30% test set. Furthermore, we normalize the portfolio weights produces by the SDF and the Moments network as in the main body of the paper: through the L^2-norm with an upper bound of $C_{\text{port}} = 1$ if the dataset is a small balanced dataset of portfolios and $C_{\text{bonds}} = 0.1$ when considering bonds.

We evaluate models the same way as for the other approaches, explained in section 3.4 of the main paper. Note here that, because of the loss of 12 months of data due to the RNN processing of sequences of macro data, the test set for the GAN is shorter than in the respective estimations of the approaches with Sharpe ratio maximization or mispricing loss minimization.

We apply this procedure with the limited hyperparameter search explained above to 3 datasets:

A. Individual bonds (see section 2 for more details)

B. Bond portfolios (see section 2 for more details)

C. 10 industry portfolios from Kenneth French's website.

With the caveat that we have not performed a more extensive architecture and hyper-parameter search, we find the following.

For the individual bonds data, we encounter to a large degree a form of *mode collapse*, even in the model with the best hyper-parameters. Mode collapse is a typical failure that can occur when training neural network architectures with a *min-max* approach, including the traditional GAN architectures for image generation. In our case, mode collapse manifests itself in the weights of the portfolio being equal to each other for each date, i.e. the portfolio produced consists of going either long or short the equally weighted portfolio in the market. Obviously such a portfolio cannot be a good approximation of the SDF. We hypothesise, without further investigation, that the mode collapse for the large dataset of bond-level data can be overcome by searching over a larger space of hyper-parameters and architectures.

For the bond portfolio data, the best model does not suffer from mode collapse. It achieves larger Sharpe, Sortino and Calmar ratios than the Sharpe ratio maximization approach and better maximum drawdown, but at a higher mean leverage than the Sharpe ratio maximization approach, and at higher typical drawdowns. In terms of alphas vs. SPY, FF3, FF5, GAN achieves alphas of the same magnitude as the Sharpe ratio maximization approach but with slightly lower significance. See Tables A.6 and A.7 for more. Table A.5 contains the hyperparameters and architectures of the best model.

For the stock portfolio data, GAN strongly outperforms the Sharpe ratio maximization approach in terms of the Sharpe ratio on test set, but with a higher maximum drawdown, and higher leverage. It performs somewhat worse than Sharpe ratio maximization in terms of excess returns against conventional risk factors, albeit neither approach provides positive and significant alphas with respect to the conventional risk factors. See Tables A.6 and A.7 for more. Table A.5 contains the hyperparameters and architectures of the best model.

Finally, training the GAN network on the same hardware as the approaches in the main paper, is a multiple of times more expensive in terms of computation hours than the networks above. More work is needed for a better understanding of the performance across different datasets of the GAN approach, as well as the two simpler approaches presented in the paper.

A.3.1 Tables - Performance of GAN relative to Sharpe ratio maximization approach

Environment		**Best model**				
Approach	Data	Architecture	bs	lr	l2reg	l1reg
GAN	Bondpf + Macro	Model 2	2048	0.001	0.01	0.01
GAN	Stockpf + Macro	Model 2	1024	0.005	none	0.01

Table A.5: Bond Portfolios - Best Models for SDF estimation

Notes: The environment describes the estimation method that was used and the data environment. The best model describes the architecture, batch size (bs), learning rate (lr) and the L^1 (l1reg) and L^2 (l2reg) regularization for the model that obtained the highest Sharpe ratio in the validation set.

Environment		**Results Test Set**						
Approach	Data	Sharpe	Sortino	Calmar	Max-DD	CAGR	Max- lev	Mean-lev
GAN	Bondpf + Macro	0.30	47.23	0.06	0.04	0.03	0.59	0.55
SR	Bondpf + Macro	0.24	12.28	0.03	0.07	0.02	1.14	0.44
GAN	Stockpf + Macro	0.14	1.85	0.01	0.52	0.07	1.91	0.95
SR	Stockpf + Macro	0.08	2.82	0.01	0.11	0.01	0.92	0.25

Table A.6: Bond and Stock Portfolios - Performance of GAN vs. Sharpe-ratio maximization approach

Notes: The environment describes the estimation method that was used and the data environment. The best model for each environment is tabulated in Table A.5 in the appendix. The results show a set of financial ratios as described in more detail in Section 3.4, as well as maximum drawdown (Max-DD), and the portfolio implied maximum leverage (Max-lev) and mean leverage (Mean-lev). Sharpe, Sortino and Calmar ratios are monthly. All units are expressed in decimal points.

Environment		**Alphas in Test Set**		
Approach	Data	S&P 500	FF 3 Factors	FF 5 Factors
GAN	Bondpf + Macro	**0.001**	0.001	0.001
		(1.96)	(1.74)	(1.59)
SR	Bondpf + Macro	**0.001**	0.001	0.001
		(2.100)	(1.900)	(1.700)
GAN	Stockpf + Macro	-0.001	-0.001	**-0.001**
		(-2.08)	(-0.67)	(-1.20)
SR	Stockpf + Macro	0.001	0.001	0.001
		(0.43)	(0.73)	(0.46)

Table A.7: Bond and Stock Portfolios - Excess Returns

Notes: The environment describes the estimation method that was used and the data environment. The best model for each environment is tabulated in Table A.5. The excess returns (alphas) are the estimate of the intercept of the SDF return series on the corresponding risk factors. The numbers in parentheses represent the t-statistic of the estimates. **Bold** means significance at the 5% level. All alphas are expressed in decimal points and represent monthly excess returns.

A.4 Summary statistics

Portfolio	median	std	25th	75th	counts
downside_1	0.477	0.980	0.102	0.932	484
downside_2	0.546	1.683	-0.108	1.244	484
downside_3	0.590	2.105	-0.293	1.484	484
downside_4	0.649	2.401	-0.577	1.786	484
downside_5	0.941	3.108	-0.639	2.191	484
idiovol_1	0.452	0.941	0.099	0.853	484
idiovol_2	0.533	1.672	-0.077	1.316	484
idiovol_3	0.607	2.075	-0.256	1.530	484
idiovol_4	0.631	2.431	-0.625	1.765	484
idiovol_5	0.884	3.126	-0.617	2.117	484
hkm_1	0.720	1.944	-0.231	1.561	484
hkm_2	0.566	1.717	-0.217	1.322	484
hkm_3	0.597	1.867	-0.156	1.389	484
hkm_4	0.587	2.011	-0.274	1.386	484
hkm_5	0.734	2.542	-0.505	1.870	484
maturity_1	0.504	1.053	0.094	1.014	484
maturity_2	0.628	1.567	-0.090	1.336	484
maturity_3	0.677	2.050	-0.371	1.609	484
maturity_4	0.676	2.166	-0.344	1.646	484
maturity_5	0.682	2.656	-0.739	2.114	484
rating_1	0.634	1.911	-0.411	1.462	484
rating_2	0.582	1.940	-0.358	1.547	484
rating_3	0.554	1.917	-0.348	1.535	484
rating_4	0.683	2.046	-0.329	1.653	484
rating_5	0.861	2.577	-0.264	1.798	484
reversal_1	0.721	1.814	-0.338	1.537	484
reversal_2	0.602	1.930	-0.404	1.570	484
reversal_3	0.585	1.893	-0.313	1.484	484
reversal_4	0.501	2.016	-0.238	1.314	484
reversal_5	0.682	2.646	-0.081	1.506	484
spread_1	0.423	1.439	-0.130	1.144	484
spread_2	0.501	1.735	-0.251	1.342	484
spread_3	0.536	1.831	-0.348	1.443	484
spread_4	0.561	1.835	-0.404	1.558	484
spread_5	0.649	1.879	-0.383	1.665	484
spread_6	0.729	1.918	-0.382	1.718	484
spread_7	0.763	1.905	-0.417	1.661	484
spread_8	0.752	1.962	-0.154	1.643	484
spread_9	0.814	2.249	-0.131	1.643	484
spread_10	1.043	3.751	-0.118	2.400	484

Table A.8: Bond Portfolio Returns - Summary Statistics

Notes: All summary statistics refer to monthly bond portfolio returns (in percent). The portfolio columns contains an acronym of the bond portfolio as described in more detail in Section 2. The sample period spans February 1977 to December 2018 which amounts to 484 monthly return observations for each portfolio.

Variable	mean	std	median	25th	75th	counts
b_treasury	0.392	0.418	0.440	0.203	0.668	17,335
b_creditspreads	-2.468	2.786	-1.631	-3.753	-0.743	17,335
rel2high	0.975	0.041	0.993	0.967	1.000	17,335
lt_rev	0.163	0.219	0.161	0.066	0.254	17,335
r12_2	0.004	0.022	0.004	-0.005	0.013	17,335
r12_7	0.017	0.056	-0.01	0.014	0.041	17,335
r2_1	0.004	0.022	0.004	-0.005	0.013	17,335
r36_13	0.076	0.146	0.065	0.016	0.135	17,335
variance40	0.016	0.023	0.008	0.005	0.017	17,335

Table A.9: Summary Statistics of features for the bond portfolios dataset

Notes: All summary statistics refer to monthly observations. Returns are expressed in percent. More details regarding the data construction can be found in Section 2.

Variable	mean	std	median	25th	75th	counts
ret_e	0.004	0.057	0.003	-0.006	0.013	438,738
coupon	6.976	2.639	7.000	5.050	8.750	438,738
lag_sp_rat	9.156	4.512	8.000	6.000	11.000	438,738
lag_moody_rat	9.118	4.183	8.000	6.000	11.000	438,738
lag_amt_out	429,445.794	551,591.483	250,000.000	100,000.000	500,000.000	438,738
lag_accrued_interest	1.857	1.292	1.613	0.800	2.675	438,738
lag_tau	11.530	9.712	8.048	4.760	17.516	438,738
lag_age	4.760	4.871	3.369	1.511	6.188	438,738
ytm	7.205	5.016	6.836	4.280	9.223	438,738
lag_bond_price	99.203	13.945	101.082	95.000	106.500	438,738
a2me	2.008	3.470	1.376	0.825	2.264	438,738
ac	-0.014	3.084	-0.001	-0.085	0.084	438,738
at	24,099.430	43,099.995	8,308.100	2,647.301	24,454.000	438,738
ato	1.945	2.708	1.643	0.996	2.591	438,738
beme	0.678	0.775	0.516	0.302	0.845	438,738
c	0.077	0.095	0.043	0.016	0.099	438,738
cf	0.059	2.488	0.053	-0.054	0.172	438,738
cf2p	0.220	0.393	0.163	0.096	0.300	438,738
cto	1.131	0.787	0.973	0.603	1.424	438,738
d2a	0.045	0.025	0.041	0.029	0.056	438,738
d2p	0.001	0.006	0.000	0.000	0.000	438,738
dpi2a	0.071	0.237	0.043	0.007	0.100	438,738
e2p	0.030	0.438	0.057	0.035	0.086	438,738
fc2y	0.226	1.264	0.181	0.101	0.288	438,738
investment	0.105	0.388	0.055	-0.006	0.131	438,738
lt_rev	0.689	1.384	0.465	0.014	1.028	438,738
lev	0.470	0.204	0.447	0.336	0.585	438,738
lturnover	1.339	1.670	0.901	0.471	1.638	438,738
ni	1,378.718	3,704.280	301.000	58.104	1,300.000	438,738
noa	0.659	0.365	0.654	0.516	0.777	438,738
oa	0.004	1.497	-0.002	-0.006	0.000	438,738
ol	0.894	0.682	0.762	0.428	1.155	438,738
op	2,961.627	5,295.489	900.000	243.326	2,965.000	438,738
pcm	0.366	0.450	0.329	0.235	0.480	438,738
pm	0.106	2.011	0.105	0.063	0.171	438,738
prof	1.368	9.862	0.769	0.504	1.163	438,738
q	0.824	0.747	0.611	0.334	1.087	438,738
rna	0.190	0.387	0.174	0.103	0.266	438,738
roa	0.055	0.064	0.056	0.027	0.085	438,738
roe	0.215	7.406	0.141	0.072	0.215	438,738
rel2high	0.855	0.169	0.912	0.785	0.990	438,738
s2p	2.085	4.059	1.202	0.588	2.334	438,738
sga2s	0.188	1.076	0.158	0.092	0.246	438,738
variance40	0.001	0.002	0.000	0.000	0.000	438,738
capm_b_mkt	0.996	0.415	0.961	0.718	1.225	438,738
ff3_ivol	0.017	0.009	0.014	0.011	0.019	438,738
lme	28,390.790	67,267.677	5,996.630	1,554.660	24,858.053	438,738
r12_2	0.116	0.351	0.093	-0.070	0.266	438,738
r12_7	0.057	0.234	0.048	-0.064	0.164	438,738
r2_1	0.011	0.099	0.010	-0.039	0.059	438,738
r36_13	0.291	0.641	0.209	-0.051	0.520	438,738

Table A.10: Summary Statistics Individual Bond Return Dataset

Notes: All summary statistics refer to monthly observations. Returns are expressed in percent. More details regarding the data construction can be found in Section 2. The sample period spans February 1973 to January 2020.

Variable	mean	std	median	25th	75th	counts
RPI	0.002	0.008	0.002	0.000	0.004	572
W875RX1	0.002	0.007	0.003	-0.000	0.005	572
DPCERA3M086SBEA	0.002	0.009	0.002	-0.000	0.005	572
CMRMTSPLx	0.002	0.011	0.002	-0.003	0.008	572
RETAILx	0.004	0.016	0.005	-0.001	0.010	572
INDPRO	0.001	0.010	0.002	-0.002	0.006	572
IPFPNSS	0.001	0.011	0.002	-0.003	0.006	572
IPFINAL	0.001	0.012	0.002	-0.003	0.007	572
IPCONGD	0.001	0.011	0.001	-0.004	0.006	572
IPDCONGD	0.001	0.035	0.001	-0.008	0.013	572
IPNCONGD	0.001	0.008	0.001	-0.004	0.006	572
IPBUSEQ	0.003	0.018	0.004	-0.004	0.010	572
IPMAT	0.002	0.010	0.002	-0.002	0.007	572
IPDMAT	0.002	0.016	0.004	-0.003	0.010	572
IPNMAT	0.001	0.011	0.001	-0.004	0.006	572
IPMANSICS	0.001	0.012	0.002	-0.002	0.007	572
IPB51222S	0.001	0.039	0.001	-0.019	0.024	572
IPFUELS	0.001	0.022	0.002	-0.011	0.011	572
CUMFNS	-0.027	0.818	0.029	-0.333	0.332	572
HWI	4.657	195.227	6.000	-97.250	103.500	572
HWIURATIO	-0.001	0.043	0.002	-0.017	0.020	572
CLF16OV	0.001	0.003	0.001	-0.000	0.002	572
CE16OV	0.001	0.007	0.001	-0.000	0.003	572
UNRATE	0.005	0.492	0.000	-0.100	0.100	572
UEMPMEAN	0.017	0.872	0.000	-0.300	0.400	572
UEMPLT5	0.000	0.102	0.002	-0.035	0.036	572
UEMP5TO14	0.001	0.092	-0.002	-0.033	0.032	572
UEMP15OV	0.004	0.063	-0.001	-0.029	0.030	572
UEMP15T26	0.004	0.093	0.000	-0.045	0.049	572
UEMP27OV	0.003	0.065	-0.003	-0.035	0.038	572
CLAIMSx	0.002	0.121	-0.003	-0.028	0.025	572
PAYEMS	0.001	0.007	0.001	0.000	0.002	572
USGOOD	-0.000	0.007	0.001	-0.002	0.002	572
CES1021000001	-0.000	0.018	0.001	-0.005	0.006	572
USCONS	0.001	0.010	0.002	-0.002	0.005	572
MANEMP	-0.001	0.006	0.000	-0.002	0.002	572
DMANEMP	-0.001	0.007	0.000	-0.002	0.002	572
NDMANEMP	-0.001	0.005	-0.000	-0.002	0.001	572
SRVPRD	0.001	0.007	0.002	0.001	0.003	572
USTPU	0.001	0.006	0.001	-0.000	0.002	572
USWTRADE	0.001	0.004	0.001	-0.000	0.002	572
USTRADE	0.001	0.008	0.001	-0.000	0.003	572
USFIRE	0.001	0.002	0.001	0.000	0.003	572
USGOVT	0.001	0.003	0.001	-0.000	0.002	572

Table A.11: Macro Series - Summary Statistics

Notes: All summary statistics refer to monthly observations. The sample period spans the period from February 1973 to September 2020

Variable	mean	std	median	25th	75th	counts
CES0600000007	40.340	0.676	40.300	39.900	40.900	572
AWOTMAN	0.000	0.145	0.000	-0.100	0.100	572
AWHMAN	40.848	0.769	40.900	40.400	41.400	572
HOUST	7.193	0.328	7.268	7.035	7.421	572
HOUSTNE	4.947	0.385	4.966	4.736	5.187	572
HOUSTMW	5.480	0.434	5.585	5.159	5.799	572
HOUSTS	6.416	0.310	6.461	6.249	6.638	572
HOUSTW	5.775	0.394	5.855	5.576	6.046	572
PERMIT	7.169	0.320	7.231	6.989	7.415	572
PERMITNE	4.975	0.355	4.973	4.762	5.193	572
PERMITMW	5.444	0.395	5.513	5.197	5.759	572
PERMITS	6.355	0.326	6.397	6.169	6.588	572
PERMITW	5.799	0.390	5.878	5.595	6.048	572
AMDMNOx	0.003	0.041	0.004	-0.017	0.028	572
ANDENOx	0.003	0.082	0.003	-0.040	0.051	572
AMDMUOx	0.004	0.010	0.003	-0.002	0.009	572
BUSINVx	0.004	0.006	0.004	0.001	0.007	572
ISRATIOx	-0.000	0.022	0.000	-0.010	0.010	572
M1SL	0.000	0.070	0.000	-0.004	0.004	572
M2SL	0.000	0.004	0.000	-0.001	0.002	572
M2REAL	0.002	0.006	0.002	-0.001	0.004	572
BOGMBASE	0.000	0.023	0.001	-0.007	0.008	572
TOTRESNS	-0.000	0.072	0.003	-0.024	0.025	572
NONBORRES	-0.000	1.141	0.003	-0.027	0.027	572
BUSLOANS	-0.000	0.009	-0.000	-0.004	0.003	572
REALLN	-0.000	0.005	0.000	-0.002	0.002	572
NONREVSL	-0.000	0.008	-0.000	-0.002	0.002	572
CONSPI	0.000	0.001	0.000	-0.000	0.001	572
S&P 500	0.006	0.037	0.009	-0.012	0.029	572
S&P: indust	0.006	0.037	0.010	-0.012	0.030	572
S&P div yield	-0.002	0.123	-0.009	-0.057	0.047	572
S&P PE ratio	0.001	0.049	0.002	-0.020	0.024	572
FEDFUNDS	-0.010	0.551	0.010	-0.080	0.100	572
CP3Mx	-0.009	0.544	0.000	-0.092	0.120	572
TB3MS	-0.009	0.458	0.000	-0.090	0.120	572
TB6MS	-0.010	0.427	0.000	-0.090	0.110	572
GS1	-0.010	0.446	0.000	-0.130	0.130	572
GS5	-0.011	0.343	-0.020	-0.180	0.160	572
GS10	-0.010	0.299	-0.020	-0.163	0.160	572
AAA	-0.008	0.241	-0.020	-0.120	0.100	572
BAA	-0.008	0.234	-0.020	-0.133	0.100	572
COMPAPFFx	0.020	0.389	0.060	-0.060	0.190	572
TB3SMFFM	-0.504	0.725	-0.250	-0.712	-0.060	572
TB6SMFFM	-0.387	0.778	-0.145	-0.590	0.030	572

Table A.12: Macro Series - Summary Statistics

Notes: All summary statistics refer to monthly observations. The sample period spans the period from February 1973 to September 2020

Variable	mean	std	median	25th	75th	counts
T1YFFM	0.004	0.790	0.120	-0.173	0.400	572
T5YFFM	0.754	1.443	0.975	0.117	1.722	572
T10YFFM	1.159	1.730	1.400	0.237	2.380	572
AAAFFM	2.312	2.010	2.520	1.278	3.750	572
BAAFFM	3.400	2.065	3.765	2.190	4.918	572
TWEXAFEGSMTHx	-0.000	0.017	0.001	-0.011	0.010	572
EXSZUSx	-0.002	0.028	-0.001	-0.019	0.016	572
EXJPUSx	-0.002	0.026	0.000	-0.016	0.015	572
EXUSUKx	-0.001	0.024	-0.000	-0.014	0.014	572
EXCAUSx	0.000	0.015	0.000	-0.008	0.010	572
WPSFD49207	-0.000	0.007	-0.000	-0.004	0.003	572
WPSFD49502	-0.000	0.010	-0.000	-0.004	0.004	572
WPSID61	0.000	0.008	-0.000	-0.003	0.003	572
WPSID62	0.000	0.046	-0.000	-0.018	0.018	572
OILPRICEx	-0.000	0.117	0.000	-0.057	0.052	572
PPICMM	-0.000	0.035	-0.001	-0.018	0.019	572
CPIAUCSL	-0.000	0.003	-0.000	-0.001	0.002	572
CPIAPPSL	-0.000	0.006	-0.000	-0.003	0.003	572
CPITRNSL	0.000	0.012	-0.000	-0.005	0.005	572
CPIMEDSL	-0.000	0.002	-0.000	-0.001	0.001	572
CUSR0000SAC	-0.000	0.006	-0.000	-0.003	0.003	572
CUSR0000SAD	0.000	0.003	-0.000	-0.002	0.002	572
CUSR0000SAS	-0.000	0.002	-0.000	-0.001	0.001	572
CPIULFSL	0.000	0.003	-0.000	-0.002	0.002	572
CUSR0000SA0L2	-0.000	0.004	-0.000	-0.002	0.002	572
CUSR0000SA0L5	-0.000	0.003	-0.000	-0.002	0.002	572
PCEPI	-0.000	0.002	0.000	-0.001	0.001	572
DDURRG3M086SBEA	-0.000	0.003	-0.000	-0.002	0.002	572
DNDGRG3M086SBEA	-0.000	0.007	-0.000	-0.003	0.003	572
DSERRG3M086SBEA	0.000	0.002	-0.000	-0.001	0.001	572
CES0600000008	-0.000	0.004	-0.000	-0.002	0.002	572
CES2000000008	-0.000	0.008	-0.000	-0.003	0.003	572
CES3000000008	-0.000	0.004	-0.000	-0.002	0.002	572
MZMSL	0.000	0.007	0.000	-0.002	0.002	572
DTCOLNVHFNM	0.000	0.027	0.000	-0.007	0.007	572
DTCTHFNM	0.000	0.022	0.000	-0.005	0.005	572
INVEST	0.000	0.010	0.000	-0.005	0.005	572
VXOCLSx	20.074	7.682	18.473	14.885	23.228	572
dp_tr_gw	-0.001	0.048	-0.004	-0.030	0.026	572
bm_tr_gw	-0.002	0.063	-0.006	-0.033	0.019	572
ntis_tr_gw	-0.000	0.008	0.000	-0.002	0.002	572
tms_tr_gw	1.663	1.266	1.755	0.690	2.670	572
dfy_tr_gw	0.001	0.125	-0.010	-0.050	0.040	572
rvar_crsp_tr_gw	0.001	2.996	-0.080	-1.633	1.520	572

Table A.13: Macro Series - Summary Statistics

Notes: All summary statistics refer to monthly observations. The sample period spans the period from February 1973 to September 2020

A.5 Predictive regressions

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
downside_1	**7.71**									0.04
	(3.58)									
downside_1	3.42	0.06	0.02	**0.05**	0.00	0.01	0.04	0.06	0.00	0.12
	(5.48)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	
downside_2	**12.63**									0.03
	(5.96)									
downside_2	1.56	0.06	0.06	**0.11**	-0.01	0.00	0.04	**0.14**	0.05	0.11
	(9.14)	(0.06)	(0.06)	(0.04)	(0.07)	(0.08)	(0.06)	(0.06)	(0.06)	
downside_3	12.00									0.02
	(7.17)									
downside_3	-1.01	0.10	0.05	**0.14**	-0.05	-0.00	0.09	0.10	-0.00	0.08
	(10.75)	(0.08)	(0.07)	(0.05)	(0.08)	(0.08)	(0.06)	(0.07)	(0.07)	
downside_4	13.20									0.02
	(7.42)									
downside_4	-1.10	0.14	0.05	**0.14**	-0.07	-0.01	0.12	0.07	-0.03	0.08
	(11.52)	(0.09)	(0.08)	(0.06)	(0.09)	(0.09)	(0.07)	(0.09)	(0.09)	
downside_5	**22.12**									0.03
	(7.58)									
downside_5	0.17	0.07	**0.19**	**0.17**	-0.04	-0.07	-0.16	0.17	0.01	0.08
	(13.31)	(0.13)	(0.09)	(0.08)	(0.10)	(0.10)	(0.12)	(0.11)	(0.12)	
idiovol_1	6.62									0.03
	(3.48)									
idiovol_1	2.36	0.05	0.02	0.04	0.00	0.01	0.04	**0.06**	-0.01	0.11
	(5.23)	(0.03)	(0.03)	(0.02)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)	
idiovol_2	11.39									0.03
	(5.92)									
idiovol_2	1.04	0.10	0.04	**0.11**	-0.02	-0.01	0.09	0.07	-0.02	0.11
	(8.94)	(0.06)	(0.05)	(0.04)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)	
idiovol_3	**14.04**									0.03
	(6.88)									
idiovol_3	2.54	0.13	0.04	**0.12**	-0.06	-0.00	0.10	0.07	-0.03	0.09
	(10.75)	(0.08)	(0.07)	(0.06)	(0.08)	(0.08)	(0.07)	(0.08)	(0.08)	
idiovol_4	14.12									0.02
	(7.44)									
idiovol_4	-0.29	0.12	0.06	**0.15**	-0.06	-0.01	0.09	0.09	-0.02	0.08
	(11.67)	(0.09)	(0.08)	(0.06)	(0.09)	(0.09)	(0.08)	(0.08)	(0.09)	
idiovol_5	**19.89**									0.02
	(7.90)									
idiovol_5	-3.20	0.01	**0.22**	**0.17**	-0.03	-0.07	-0.20	**0.22**	0.14	0.09
	(13.28)	(0.11)	(0.09)	(0.07)	(0.09)	(0.10)	(0.12)	(0.11)	(0.13)	
hkm_1	**14.00**									0.03
	(4.58)									
hkm_1	5.73	0.04	0.06	0.10	0.01	-0.01	-0.01	0.04	-0.03	0.06
	(9.12)	(0.07)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.07)	(0.08)	
hkm_2	**13.33**									0.03
	(5.72)									
hkm_2	4.98	0.07	0.03	**0.11**	-0.01	0.02	0.07	0.09	-0.02	0.10
	(8.69)	(0.07)	(0.06)	(0.04)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	
hkm_3	12.96									0.03
	(6.74)									
hkm_3	4.21	0.11	0.02	**0.09**	-0.03	0.01	0.10	0.08	-0.02	0.09
	(9.79)	(0.07)	(0.06)	(0.05)	(0.08)	(0.08)	(0.06)	(0.06)	(0.06)	
hkm_4	12.21									0.02
	(6.96)									

Table A.14: In-sample predictive Regression Table 1

Notes: The table reports estimates from OLS regressions of excess bond returns on the lagged variables. The dependent variable rx_{t+1}, is the excess monthly log return on the corporate bond portfolio. CP08 is the weighted combination of forward spreads on the treasury yield curve as constructed in Cochrane and Piazessi (2008). PC1-PC8 denote factors estimated by principal components analysis using a panel of data with 132 individual series over the period 1977:01-2018:11. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The sample period spans 1977:02-2018:12

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
hkm_4	1.22	0.10	0.04	**0.13**	-0.04	-0.00	0.07	0.08	0.01	0.09
	(10.19)	(0.06)	(0.06)	(0.05)	(0.08)	(0.09)	(0.06)	(0.06)	(0.06)	
hkm_5	**17.18**									0.03
	(7.22)									
hkm_5	-10.50	0.09	**0.21**	**0.18**	-0.07	-0.09	-0.04	0.18	0.02	0.10
	(11.75)	(0.10)	(0.08)	(0.07)	(0.09)	(0.09)	(0.08)	(0.09)	(0.09)	
maturity_1	6.18									0.02
	(3.71)									
maturity_1	-0.74	0.07	0.04	0.05	-0.01	-0.02	0.02	0.07	-0.01	0.09
	(5.55)	(0.04)	(0.04)	(0.03)	(0.04)	(0.05)	(0.04)	(0.04)	(0.04)	
maturity_2	**11.60**									0.03
	(5.39)									
maturity_2	0.63	0.08	0.06	**0.09**	-0.02	-0.02	0.04	0.08	-0.01	0.09
	(8.14)	(0.06)	(0.05)	(0.04)	(0.06)	(0.07)	(0.05)	(0.05)	(0.05)	
maturity_3	**15.42**									0.03
	(6.24)									
maturity_3	-1.98	0.07	0.12	**0.13**	-0.03	-0.04	0.03	0.12	-0.01	0.09
	(10.15)	(0.09)	(0.06)	(0.06)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	
maturity_4	**15.43**									0.03
	(6.89)									
maturity_4	1.80	0.09	0.05	**0.14**	-0.03	-0.01	0.08	0.09	-0.01	0.09
	(10.68)	(0.08)	(0.07)	(0.05)	(0.08)	(0.09)	(0.06)	(0.07)	(0.07)	
maturity_5	**16.34**									0.02
	(7.36)									
maturity_5	-1.53	0.13	0.04	**0.22**	-0.06	-0.01	**0.14**	0.11	-0.03	0.11
	(12.30)	(0.09)	(0.08)	(0.07)	(0.08)	(0.09)	(0.07)	(0.10)	(0.11)	
rating_1	12.18									0.02
	(6.26)									
rating_1	1.85	**0.12**	-0.00	**0.14**	-0.06	0.01	**0.14**	0.08	-0.03	0.12
	(9.43)	(0.05)	(0.06)	(0.05)	(0.07)	(0.07)	(0.05)	(0.06)	(0.07)	
rating_2	**12.39**									0.02
	(6.17)									
rating_2	0.03	0.11	0.03	**0.15**	-0.03	-0.00	0.11	0.07	-0.04	0.11
	(9.52)	(0.07)	(0.06)	(0.05)	(0.07)	(0.08)	(0.05)	(0.06)	(0.06)	
rating_3	**13.45**									0.03
	(6.14)									
rating_3	2.70	0.12	0.04	**0.10**	-0.05	-0.00	0.06	0.08	-0.04	0.09
	(9.36)	(0.08)	(0.06)	(0.04)	(0.07)	(0.08)	(0.06)	(0.06)	(0.06)	
rating_4	**12.66**									0.02
	(6.35)									
rating_4	-2.30	0.07	0.10	0.10	-0.01	-0.03	0.04	0.11	0.02	0.07
	(10.07)	(0.09)	(0.07)	(0.06)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	
rating_5	**18.19**									0.03
	(5.65)									
rating_5	-3.75	-0.00	**0.22**	0.12	0.02	-0.10	-0.14	0.15	0.04	0.08
	(11.75)	(0.12)	(0.07)	(0.08)	(0.08)	(0.08)	(0.10)	(0.09)	(0.09)	
reversal_1	**12.58**									0.03
	(5.30)									
reversal_1	2.72	0.08	0.06	**0.11**	-0.01	0.01	0.04	0.07	-0.00	0.09
	(8.88)	(0.07)	(0.06)	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	
reversal_2	**12.13**									0.02
	(5.83)									
reversal_2	3.84	0.07	0.02	**0.11**	-0.01	0.00	0.07	0.05	-0.02	0.07

Table A.15: In-sample predictive Regression Table 2

Notes: The table reports estimates from OLS regressions of excess bond returns on the lagged variables. The dependent variable rx_{t+1}, is the excess monthly log return on the corporate bond portfolio. CP08 is the a weighted combination of forward spreads on the treasury yield curve as constructed in Cochrane and Piazzessi (2008). PC1-PC8 denote factors estimated by principal components analysis using a panel of data with 132 individual series over the period 1977:01-2018:11. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The sample period spans 1977:02-2018:12

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
	(9.62)	(0.07)	(0.06)	(0.05)	(0.07)	(0.07)	(0.06)	(0.06)	(0.07)	
reversal_3	**13.05**									0.03
	(6.03)									
reversal_3	0.49	0.10	0.04	**0.13**	-0.03	-0.01	0.10	**0.13**	0.00	0.11
	(9.82)	(0.06)	(0.06)	(0.05)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	
reversal_4	11.10									0.02
	(6.78)									
reversal_4	-0.43	**0.15**	0.03	**0.12**	-0.06	-0.02	0.09	0.04	-0.03	0.09
	(10.40)	(0.06)	(0.06)	(0.05)	(0.08)	(0.08)	(0.06)	(0.07)	(0.06)	
reversal_5	**21.26**									0.04
	(7.36)									
reversal_5	0.18	0.09	0.13	**0.15**	-0.07	-0.06	-0.08	**0.27**	0.01	0.10
	(11.44)	(0.10)	(0.08)	(0.07)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	
spread_1	**9.53**									0.02
	(4.11)									
spread_1	-2.68	**0.10**	0.02	**0.12**	-0.03	-0.04	**0.10**	0.07	-0.03	0.12
	(6.92)	(0.04)	(0.04)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	
spread_2	**10.39**									0.02
	(5.00)									
spread_2	-2.28	0.10	0.02	**0.14**	-0.04	-0.00	**0.13**	0.10	-0.06	0.13
	(8.04)	(0.06)	(0.05)	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	
spread_3	**11.82**									0.02
	(5.04)									
spread_3	0.46	0.10	0.02	**0.14**	-0.03	-0.00	**0.10**	0.08	-0.04	0.11
	(8.38)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)	(0.07)	
spread_4	**11.04**									0.02
	(4.78)									
spread_4	1.38	0.09	0.02	**0.13**	-0.02	0.01	0.10	0.06	-0.05	0.09
	(8.53)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.07)	(0.07)	
spread_5	10.27									0.02
	(5.38)									
spread_5	0.42	0.11	0.03	**0.12**	-0.03	0.02	0.09	0.07	-0.04	0.09
	(8.49)	(0.06)	(0.05)	(0.04)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	
spread_6	10.72									0.02
	(5.54)									
spread_6	0.45	0.09	0.05	**0.10**	-0.02	0.01	0.08	0.08	-0.03	0.08
	(8.63)	(0.06)	(0.06)	(0.04)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	
spread_7	**11.12**									0.02
	(5.17)									
spread_7	-2.45	0.07	0.09	**0.10**	-0.01	-0.02	0.05	0.11	-0.02	0.07
	(8.82)	(0.08)	(0.06)	(0.05)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	
spread_8	10.37									0.02
	(5.48)									
spread_8	-3.00	0.04	0.12	0.09	0.01	-0.02	0.02	0.10	0.05	0.06
	(9.17)	(0.08)	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	
spread_9	**11.55**									0.01
	(5.54)									
spread_9	-8.91	0.02	**0.17**	0.12	0.01	-0.08	-0.02	0.13	0.05	0.06
	(10.27)	(0.10)	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	
spread_10	**23.56**									0.02
	(6.12)									
spread_10	-7.76	-0.08	**0.34**	0.13	0.05	-0.16	**-0.36**	**0.30**	-0.00	0.10
	(14.61)	(0.17)	(0.11)	(0.12)	(0.11)	(0.10)	(0.15)	(0.15)	(0.14)	

Table A.16: In-sample predictive Regression Table 3

Notes: The table reports estimates from OLS regressions of excess returns of bond portfolios on the lagged variables. The dependent variable rx_{t+1} is the excess monthly log return on the corporate bond portfolio. CP08 is the a weighted combination of forward spreads on the treasury yield curve as constructed in Cochrane and Piazessi (2008). PC1-PC8 denote factors estimated by principal components analysis using a panel of data with 132 individual series over the period 1977:01-2018:11. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The sample period spans 1977:02-2018:12

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
downside_1	**6.55**									0.04
	(2.84)									
downside_1	**6.93**	**0.05**	0.01	0.03	0.00	0.03	0.01	0.00	-0.03	0.10
	(2.67)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
downside_2	**11.38**									0.04
	(4.70)									
downside_2	**10.46**	**0.07**	0.00	**0.06**	-0.01	0.06	0.05	-0.01	-0.05	0.10
	(4.53)	(0.03)	(0.06)	(0.03)	(0.05)	(0.05)	(0.05)	(0.06)	(0.05)	
downside_3	**11.47**									0.02
	(5.55)									
downside_3	10.67	**0.08**	0.01	0.06	0.00	0.10	0.03	-0.02	-0.08	0.07
	(5.54)	(0.03)	(0.07)	(0.03)	(0.05)	(0.07)	(0.05)	(0.06)	(0.06)	
downside_4	**12.10**									0.02
	(5.78)									
downside_4	**11.42**	**0.09**	0.00	0.06	0.02	0.10	0.01	-0.03	-0.07	0.06
	(5.82)	(0.04)	(0.08)	(0.04)	(0.06)	(0.07)	(0.07)	(0.08)	(0.07)	
downside_5	**18.74**									0.03
	(6.13)									
downside_5	12.04	0.09	-0.12	0.03	-0.02	0.11	-0.05	-0.13	-0.16	0.09
	(6.51)	(0.06)	(0.09)	(0.06)	(0.07)	(0.10)	(0.10)	(0.12)	(0.09)	
idiovol_1	**5.48**									0.03
	(2.70)									
idiovol_1	**5.91**	**0.05**	0.01	0.03	-0.00	0.03	0.00	0.01	-0.03	0.10
	(2.47)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	
idiovol_2	**9.82**									0.03
	(4.72)									
idiovol_2	**9.16**	**0.08**	0.00	**0.05**	0.01	0.07	0.02	-0.02	-0.05	0.09
	(4.56)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.05)	
idiovol_3	**12.96**									0.03
	(5.36)									
idiovol_3	**13.11**	**0.09**	0.02	0.05	0.01	0.08	-0.01	-0.03	-0.09	0.08
	(5.33)	(0.04)	(0.07)	(0.03)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	
idiovol_4	**13.26**									0.02
	(5.80)									
idiovol_4	**12.15**	**0.09**	0.00	0.06	0.01	0.11	0.02	-0.03	-0.08	0.07
	(5.83)	(0.04)	(0.08)	(0.04)	(0.06)	(0.08)	(0.06)	(0.08)	(0.07)	
idiovol_5	**17.15**									0.02
	(6.39)									
idiovol_5	9.64	0.07	-0.12	0.03	-0.02	0.10	0.04	-0.12	-0.12	0.07
	(7.00)	(0.05)	(0.09)	(0.06)	(0.07)	(0.09)	(0.12)	(0.14)	(0.09)	
hkm_1	**10.74**									0.03
	(3.61)									
hkm_1	5.94	0.04	-0.09	0.03	0.03	0.07	0.05	-0.03	-0.08	0.07
	(3.83)	(0.03)	(0.06)	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	
hkm_2	**11.63**									0.04
	(4.51)									
hkm_2	**10.38**	**0.06**	-0.01	0.05	0.00	0.08	0.03	-0.00	-0.06	0.09
	(4.42)	(0.03)	(0.06)	(0.03)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)	
hkm_3	**12.38**									0.04
	(5.24)									
hkm_3	**13.41**	**0.08**	0.04	0.05	0.01	0.07	0.02	-0.01	-0.07	0.08
	(5.03)	(0.03)	(0.07)	(0.03)	(0.05)	(0.06)	(0.05)	(0.05)	(0.05)	
hkm_4	**12.92**									0.03
	(5.33)									

Table A.17: Out-of-sample predictive Regression Table 1

Notes: The table reports estimates from OLS regressions of excess bond returns on the lagged variables. The dependent variable rx_{t+1}, is the excess monthly log return on the corporate bond portfolio. All factors are constructed with information in the t information set only. In particular, the Cochrane-Piazessi factor, CP08, is recursively estimated in analogy to Cochrane and Piazessi (2008). PC1-PC8 denote factors estimated recursively by principal components analysis using a panel of data with 132 individual series. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The full sample period spans 1977:02-2018:12.

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
hkm_4	**12.41**	**0.08**	0.04	**0.07**	-0.03	0.08	0.04	-0.04	-0.06	0.09
	(5.12)	(0.03)	(0.07)	(0.03)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)	
hkm_5	**15.81**									0.03
	(5.59)									
hkm_5	10.47	**0.10**	-0.10	0.04	-0.00	0.12	0.05	-0.02	-0.10	0.10
	(5.72)	(0.04)	(0.08)	(0.04)	(0.06)	(0.08)	(0.07)	(0.08)	(0.07)	
maturity_1	5.05									0.02
	(2.83)									
maturity_1	5.11	**0.06**	-0.00	0.02	-0.00	0.04	-0.01	0.00	-0.04	0.08
	(2.70)	(0.02)	(0.04)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	
maturity_2	**9.53**									0.03
	(4.17)									
maturity_2	**8.49**	**0.07**	-0.01	0.03	0.01	0.07	0.02	-0.00	-0.05	0.08
	(4.03)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.04)	(0.05)	(0.04)	
maturity_3	**13.02**									0.03
	(4.97)									
maturity_3	**10.31**	**0.08**	-0.05	0.03	0.00	0.10	0.02	0.01	-0.10	0.08
	(4.97)	(0.04)	(0.07)	(0.04)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	
maturity_4	**13.77**									0.03
	(5.34)									
maturity_4	**11.88**	**0.08**	-0.02	0.05	0.02	0.09	0.05	-0.01	-0.08	0.08
	(5.27)	(0.04)	(0.08)	(0.04)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	
maturity_5	**15.67**									0.03
	(5.68)									
maturity_5	**12.08**	0.09	-0.05	**0.10**	0.01	0.09	0.08	-0.01	-0.10	0.09
	(5.80)	(0.05)	(0.09)	(0.04)	(0.06)	(0.07)	(0.07)	(0.08)	(0.09)	
rating_1	**12.11**									0.03
	(4.72)									
rating_1	**11.94**	**0.08**	0.02	**0.07**	0.01	0.05	0.06	-0.02	-0.05	0.09
	(4.69)	(0.03)	(0.07)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	
rating_2	**11.62**									0.03
	(4.84)									
rating_2	**9.93**	**0.08**	-0.01	**0.07**	0.00	0.08	0.04	-0.02	-0.06	0.09
	(4.73)	(0.03)	(0.07)	(0.03)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	
rating_3	**11.87**									0.03
	(4.80)									
rating_3	**11.27**	**0.09**	-0.00	0.04	0.01	0.08	0.02	0.01	-0.06	0.08
	(4.66)	(0.03)	(0.06)	(0.03)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	
rating_4	**10.78**									0.02
	(5.00)									
rating_4	8.80	**0.08**	-0.02	0.02	0.01	0.11	0.05	0.01	-0.11	0.08
	(4.95)	(0.04)	(0.07)	(0.04)	(0.05)	(0.07)	(0.06)	(0.07)	(0.06)	
rating_5	**13.40**									0.02
	(4.60)									
rating_5	6.93	0.05	**-0.16**	-0.00	0.01	0.11	-0.02	0.01	-0.14	0.07
	(5.20)	(0.05)	(0.07)	(0.05)	(0.06)	(0.09)	(0.09)	(0.09)	(0.08)	
reversal_1	**12.53**									0.04
	(4.24)									
reversal_1	**10.66**	**0.07**	-0.03	0.04	-0.02	0.06	0.03	0.01	-0.02	0.09
	(4.07)	(0.03)	(0.05)	(0.03)	(0.04)	(0.06)	(0.05)	(0.06)	(0.06)	
reversal_2	**11.44**									0.03
	(4.59)									
reversal_2	**10.75**	0.06	0.01	0.05	0.01	0.07	0.02	-0.06	-0.03	0.06

Table A.18: Out-of-sample predictive Regression Table 2

Notes: The table reports estimates from OLS regressions of excess bond returns on the lagged variables. The dependent variable rx_{t+1}, is the excess monthly log return on the corporate bond portfolio. All factors are constructed with information in the t information set only. In particular, the Cochrane-Piazessi factor, CP08, is recursively estimated in analogy to Cochrane and Piazessi (2008). PC1-PC8 denote factors estimated recursively by principal components analysis using a panel of data with 132 individual series. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The full sample period spans 1977:02-2018:12.

PF	CP08	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	R^2
	(4.60)	(0.03)	(0.06)	(0.03)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	
reversal_3	**11.57**									0.03
	(4.71)									
reversal_3	**11.65**	**0.08**	0.02	**0.06**	0.01	0.06	0.06	0.01	-0.05	0.09
	(4.73)	(0.03)	(0.06)	(0.03)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	
reversal_4	9.13									0.02
	(5.29)									
reversal_4	8.18	**0.08**	0.01	0.06	0.01	0.07	0.03	-0.05	-0.06	0.07
	(5.34)	(0.03)	(0.07)	(0.03)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	
reversal_5	**17.01**									0.03
	(5.77)									
reversal_5	**15.19**	**0.10**	-0.06	0.06	-0.04	0.06	-0.10	-0.02	**-0.21**	0.10
	(6.10)	(0.05)	(0.08)	(0.05)	(0.06)	(0.08)	(0.08)	(0.10)	(0.09)	
spread_1	**8.93**									0.03
	(3.37)									
spread_1	**9.01**	**0.06**	-0.01	**0.07**	0.02	0.01	0.03	0.01	-0.01	0.10
	(3.42)	(0.02)	(0.05)	(0.02)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	
spread_2	**9.96**									0.03
	(3.78)									
spread_2	**9.04**	**0.07**	-0.01	**0.07**	0.01	0.05	0.06	0.00	-0.03	0.09
	(3.97)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	
spread_3	**11.49**									0.03
	(3.89)									
spread_3	**9.68**	**0.07**	-0.03	**0.07**	0.00	0.06	0.05	0.01	-0.06	0.10
	(4.08)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	
spread_4	**10.42**									0.03
	(3.72)									
spread_4	**8.72**	**0.07**	-0.02	0.06	0.02	0.06	0.05	-0.00	-0.05	0.08
	(3.95)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	(0.06)	
spread_5	**9.07**									0.02
	(4.12)									
spread_5	7.52	**0.08**	-0.01	0.05	-0.00	0.07	0.05	0.00	-0.02	0.07
	(4.12)	(0.03)	(0.06)	(0.03)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	
spread_6	**9.34**									0.02
	(4.34)									
spread_6	**8.56**	**0.08**	-0.01	0.04	0.04	0.08	0.05	0.01	-0.04	0.07
	(4.34)	(0.03)	(0.06)	(0.03)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	
spread_7	**9.28**									0.02
	(4.04)									
spread_7	7.80	**0.08**	-0.03	0.02	0.02	0.09	0.03	0.03	-0.08	0.07
	(4.11)	(0.03)	(0.06)	(0.03)	(0.04)	(0.06)	(0.05)	(0.06)	(0.05)	
spread_8	**9.19**									0.02
	(4.27)									
spread_8	6.99	**0.07**	-0.03	0.01	0.02	0.12	0.04	0.01	-0.09	0.07
	(4.38)	(0.04)	(0.06)	(0.04)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	
spread_9	**9.05**									0.01
	(4.42)									
spread_9	4.99	0.06	-0.09	-0.00	0.03	0.11	0.03	0.02	-0.11	0.06
	(4.88)	(0.04)	(0.07)	(0.04)	(0.05)	(0.07)	(0.07)	(0.08)	(0.07)	
spread_10	**15.22**									0.01
	(5.02)									
spread_10	3.26	0.04	**-0.27**	-0.05	-0.03	0.15	-0.11	-0.08	**-0.29**	0.09
	(6.86)	(0.08)	(0.10)	(0.08)	(0.10)	(0.13)	(0.14)	(0.14)	(0.11)	

Table A.19: Out-of-sample predictive Regression Table 3

Notes: The table reports estimates from OLS regressions of excess bond returns on the lagged variables. The dependent variable rx_{t+1}, is the excess monthly log return on the corporate bond portfolio. All factors are constructed with information in the t information set only. In particular, the Cochrane-Piazessi factor, CP08, is recursively estimated in analogy to Cochrane and Piazessi (2008). PC1-PC8 denote factors estimated recursively by principal components analysis using a panel of data with 132 individual series. Heteroskedasticity robust standard errors are reported in parentheses. Coefficients that are statistically significant at the 5% or better level are highlighted in bold. All specifications include a constant. The full sample period spans 1977:02-2018:12

A.6 Description of fundamentals

Acronym	Name	Definition	Reference
at	Total Assets	Total Assets (at)	Gandhi and Lustig (2015)
ato	Net sales over lagged	Net sales (sale) over lagged net operating assets. Net operating assets are the difference between operating assets and operating liabilities (defined in noa)	Soliman (2008)
beme	Book to Market Ratio	Book equity is shareholder equity (sh) plus deferred taxes and investment tax credit (txditc), minus preferred stock (ps). "sh" is shareholders' equity (seq). If missing, "sh" is the sum of common equity (ceq) and preferred stock (ps). If missing, "sh" is the difference between total assets (at) and total liabilities (lt). Depending on availability, we use the redemption (item pstkrv), liquidating (item pstkl), or par value (item pstk) for ps. The market value of equity (prc*shrout) is as of December t-1.	Fama and French (1992)
b_treasury	Treasuries Beta	Coefficient of the bond portfolio excess return from the regression on the average excess returns (across maturities) in the treasury market.	
b_creditspread	Creditspread Beta	Coefficient of the bond portfolio excess return from the regression on the credit spread (BAA-Treasury).	
c	Ratio of cash and short-term investments to total assets	Ratio of cash and short-term investments (che) to total assets (at)	Palazzo (2012)
cf	Free Cash Flow to Book Value	Cash flow to book value of equity is the ratio of net income (ni), depreciation and amortization (DP), less change in working capital (wcapch), and capital expenditure (capx) over the book-value of equity (defined in beme)	Hou et al. (2011)
cf2p	Cash flow to price	Cash flow over market capitalization (prc*shrout) as of December t-1. Cashow is defined as income before extraordinary items (ib) plus depreciation and amortization (dp) plus deferred taxes (txdb).	Desai et al. (2004)
cto	Capital turnover	Ratio of net sales (sale) to lagged total assets (at)	Haugen and Baker (1996)
d2a	Capital intensity	Ratio of depreciation and amortization (dp) to total assets (at)	Gorodnichenko and Weber (2016)
d2p	Dividend Yield	Total dividends (divamt) paid from July of t-1 to June of t per dollar of equity (LME) in June of t	Litzenberger and Ramaswamy (1979)
dp12a	Change in property, plants, and equipment	Changes in property, plants, and equipment (ppegt) and inventory (invt) over lagged total assets (at)	Lyandres et al. (2008)
e2p	Earnings to price	The earnings used in June of year t are total earnings before extraordinary items for the last fiscal year end in t-1. P (actually ME) is price times shares outstanding at the end of December of t-1.	Basu (1983)

Table A.20: Fundamentals - Description

Acronym	Name	Definition	Reference
fc2y	Fixed costs to sales	Ratio of selling, general, and administrative expenses (xsgs), research and development expenses (xrd), and advertising expenses (xad) to net sales (sale)	D'Acunto et al. (2018)
investment	Investment	Change in total assets (at) from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets	Cooper et al. (2008)
lme	Size	Total market capitalization at the end of the previous month defined as price times shares outstanding	Fama and French (1992)
lt_rev	Long-term reversal	Cumulative return from 60 months before the return prediction to 13 months before	Jegadeesh and Titman (1993)
lev	Leverage	Ratio of long-term debt (dltt) and debt in current liabilities (dlc) to the sum of long-term debt, debt in current liabilities, and stockholders' equity (seq)	Lewellen (2015)
lturnover	Turnover	Turnover is last month's volume (vol) over shares outstanding (shrout)	Datar et al. (1998)
ni	Net Share Issues	The change in the natural log of split-adjusted shares outstanding (csho*ajex) from the fiscal yearend in t-2 to the fiscal yearend in t-1	Pontiff and Woodgate (2008)
noa	Net operating assets	Difference between operating assets minus operating liabilities scaled by lagged total assets (at). Operating assets are total assets (at) minus cash and short-term investments (che), minus investment and other advances (ivao). Operating liabilities are total assets (at), minus debt in current liabilities (dlc), minus long-term debt (dltt), minus minority interest (mib), minus preferred stock (pstk), minus common equity (ceq).	Hirshleifer et al. (2004)
oa	Operating accruals	Changes in non-cash working capital minus depreciation (dp) scaled by lagged total assets (at). Non-cash working capital is defined in Accrual (ac)	Sloan (1996)
ol	Operating leverage	Sum of cost of goods sold (cogs) and selling, general, and administrative expenses (xsga) over total assets (at)	Novy-Marx (2011)
op	Operating profitability	Annual revenues (revt) minus cost of goods sold (cogs), interest expense (tie), and selling, general, and administrative expenses (xsga) divided by book equity (defined in beme)	Fama and French (2015b)
pcm	Price to cost margin	Difference between net sales (sale) and costs of goods sold (cogs) divided by net sales (sale)	Bustamante and Donangelo (2017)

Table A.21: Fundamentals - Description cont'd

Acronym	Name	Definition	Reference
pm	Profit margin	Operating income after depreciation (oiadp) over net sales (sale)	Soliman (2008)
prof	Profitability	Gross profitability (gp) divided by the book value of equity (defined in beme)	Ball et al. (2015)
q	Tobin's Q	Tobin's Q is total assets (at), the market value of equity (prc*shrout) minus cash and short-term investments (che), minus deferred taxes (txdb) scaled by total assets (at)	Kaldor (1966)
rna	Return on net operating assets	Ratio of operating income after depreciation (oiadp) to lagged net operating assets.	Soliman (2008)
roa	Return on assets	Income before extraordinary items (ib) to lagged total assets (at)	Balakrishnan et al. (2010)
roe	Return on equity	Income before extraordinary items (ib) to lagged book-value of equity	Haugen and Baker (1996)
rel2high	Closeness to past year high	The ratio of stock price at the end of the previous calendar month and the highest daily price in the past year	George and Hwang (2004)
s2p	Sales to price	Ratio of net sales (sale) to the market capitalization (lme)	Lewellen (2015)
sga2s	Selling, general and administrative expenses to sales	Ratio of selling, general and administrative expenses (xsga) to net sales (sale)	Freyberger et al. (2020a)
spread	Bid-ask spread	The average daily bid-ask spread in the previous month	Chung and Zhang (2014)
variance40	Variance	Variance of daily returns in the past 40 trading days	Ang et al. (2006)
capm_b_mkt	Market Beta	Coefficient of the market excess return from the regression on excess returns. The estimation window is 252 trading days, and the minimum window is 126 trading days	Fama and MacBeth (1973)
ff3_ivol	Idiosyncratic volatility	Standard deviation of the residuals from a regression of excess returns on the Fama and French three-factor model. The estimation window is 252 trading days, and the minimum window is 126 trading days	Ang et al. (2006)
r12_2	Momentum	Cumulative return from 12 months before the return prediction to two months before	Fama et al. (1996)
r12_7	Intermediate momentum	Cumulative return from 12 months before the return prediction to seven months before	Novy-Marx (2012)
r2_1	Short-term momentum	Lagged one-month return	Jegadeesh and Titman (1993)
r36_13	Long-term momentum	Cumulative return from 36 months before the return prediction to 13 months before	De Bondt and Thaler (1985)

Table A.22: Fundamentals - Description cont'd

Acronym	Name	Definition	Reference
dp_tr_gw	Dividend-price ratio	We use the monthly CRSP market return index from December 1971 to December 2020 and the risk-free rate from Kenneth French website. The aggregate dividend is obtained from the difference between cum (vwretd) and ex dividend (vwretx) returns. Then we take a rolling sum of 12 months, and construct the dp variable as log of the rolling sum minus log of the totalreturn cum dividend. Stationarity transformation in the end: first differences	Koijen and Van Nieuwerburgh (2011)
bm_tr_gw	Book-to-Market Ratio	From December 1971 to December 2019 we use the bm variable from Welch and Goyal (2008). Note that these are based on the Dow Jones Industrial Average. For the remaining months of 2020, we use the price/book ratio from the last-day-of month from Bloomberg (we also invert it to get a book/price value). Stationarity transformation in the end: first differences of logs.	Welch and Goyal (2008)
ntis_tr_gw	Net Equity Expansion	We use all of CRSP. We use formula (3) in Welch and Goyal (2008). We take the TOTVAL variable (monthly) from CRSP as market cap variable, and we used VWRETX to construct net issuance. Then take 12-month moving sum, and divide by end-of-year market cap. Stationarity transformation in the end: first differences	Welch and Goyal (2008)
tms_tr_gw	Term spread	The term spread is the difference between long-term yield on govt bonds (lty) and the treasury bill (tbl); we use GS10 in FRED for lty, so that we calculate tms from FRED as GS10-TB3MS. Stationarity transformation in the end: no transformation	Welch and Goyal (2008)
dfy_tr_gw	Default spread	The difference between BAA and AAA-rated corporate bond yields; corporate bond yields both for BAA and AAA are from FRED. Stationarity transformation in the end: first differences	Welch and Goyal (2008)
svar_tr_gw	Stock variance	We construct the monthly variance as the sum of squared returns within a month of the market return from CRSP. Stationarity transformation in the end: difference of logs	Welch and Goyal (2008)
a2me	Assets to market cap	Total assets (at) over market capitalization (prc*shrout) as of December t-1	Bhandari (1988)
ac	Accrual	Change in operating working capital per split-adjusted share from the fiscal year end t-2 to t-1 divided by book equity (defined in beme) per share in t-1. Operating working capital per split-adjusted share is defined as current assets (act) minus cash and short-term investments (che) minus current liabilities (lct) minus debt in current liabilities (dlc) minus income taxes payable (txp).	Sloan (1996)

Table A.23: Fundamentals - Description cont'd

A.6.1 Description of the macro variables

Acronym	Definition	tcode
RPI	Real Personal Income	5
W875RX1	Real personal income ex transfer receipts	5
INDPRO	IP Index	5
IPFPNSS	IP: Final Products and Nonindustrial Supplies	5
IPFINAL	IP: Final Products (Market Group)	5
IPCONGD	IP: Consumer Goods	5
IPDCONGD	IP: Durable Consumer Goods	5
IPNCONGD	IP: Nondurable Consumer Goods	5
IPBUSEQ	IP: Business Equipment	5
IPMAT	IP: Materials	5
IPDMAT	IP: Durable Materials	5
IPNMAT	IP: Nondurable Materials	5
IPMANSICS	IP: Manufacturing (SIC)	5
IPB51222S	IP: Residential Utilities	5
IPFUELS	IP: Fuels	5
CUMFNS	Capacity Utilization: Manufacturing	2
HWI	Help-Wanted Index for United States	2
HWIURATIO	Ratio of Help Wanted/No. Unemployed	2
CLF16OV	Civilian Labor Force	5
CE16OV	Civilian Employment	5
UNRATE	Civilian Unemployment Rate	2
UEMPMEAN	Average Duration of Unemployment (Weeks)	2
UEMPLT5	Civilians Unemployed - Less Than 5 Weeks	5
UEMP5TO14	Civilians Unemployed for 5-14 Weeks	5
UEMP15OV	Civilians Unemployed - 15 Weeks & Over	5
UEMP15T26	Civilians Unemployed for 15-26 Weeks	5
UEMP27OV	Civilians Unemployed for 27 Weeks and Over	5
CLAIMSx	Initial Claims	5
PAYEMS	All Employees: Total nonfarm	5
USGOOD	All Employees: Goods-Producing Industries	5
CES1021000001	All Employees: Mining and Logging: Mining	5
USCONS	All Employees: Construction	5
MANEMP	All Employees: Manufacturing	5
DMANEMP	All Employees: Durable goods	5
NDMANEMP	All Employees: Nondurable goods	5
SRVPRD	All Employees: Service-Providing Industries	5
USTPU	All Employees: Trade, Transportation & Utilities	5
USWTRADE	All Employees: Wholesale Trade	5
USTRADE	All Employees: Retail Trade	5
USFIRE	All Employees: Financial Activities	5
USGOVT	All Employees: Government	5
CES0600000007	Avg Weekly Hours : Goods-Producing	1
AWOTMAN	Avg Weekly Overtime Hours : Manufacturing	2
AWHMAN	Avg Weekly Hours : Manufacturing	1
CES0600000008	Avg Hourly Earnings : Goods-Producing	6
CES2000000008	Avg Hourly Earnings : Construction	6
CES3000000008	Avg Hourly Earnings : Manufacturing	6
HOUST	Housing Starts: Total New Privately Owned	4
HOUSTNE	Housing Starts, Northeast	4

Table A.24: Macro Variables - Description

Notes: The table tabulates the variable acronym of the macro time-series, the description, as well as, the transformation code ("tcode"). The transformation code denotes the variable transformation to make the series covariance stationary, (McCracken and Ng, 2016). The tcode corresponds to: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$.

Acronym	Definition	tcode
HOUSTMW	Housing Starts, Midwest	4
HOUSTS	Housing Starts, South	4
HOUSTW	Housing Starts, West	4
PERMIT	New Private Housing Permits (SAAR)	4
PERMITNE	New Private Housing Permits, Northeast (SAAR)	4
PERMITMW	New Private Housing Permits, Midwest (SAAR)	4
PERMITS	New Private Housing Permits, South (SAAR)	4
PERMITW	New Private Housing Permits, West (SAAR)	4
DPCERA3M086SBEA	Real personal consumption expenditures	5
CMRMTSPLx	Real Manu. and Trade Industries Sales	5
RETAILx	Retail and Food Services Sales	5
AMDMNOx	New Orders for Durable Goods	5
ANDENOx	New Orders for Nondefense Capital Goods	5
AMDMUOx	Unfilled Orders for Durable Goods	5
BUSINVx	Total Business Inventories	5
ISRATIOx	Total Business: Inventories to Sales Ratio	2
M1SL	M1 Money Stock	6
M2SL	M2 Money Stock	6
M2REAL	Real M2 Money Stock	5
BOGMBASE	Monetary Base	6
TOTRESNS	Total Reserves of Depository Institutions	6
NONBORRES	Reserves Of Depository Institutions	7
BUSLOANS	Commercial and Industrial Loans	6
REALLN	Real Estate Loans at All Commercial Banks	6
NONREVSL	Total Nonrevolving Credit	6
CONSPI	Nonrevolving consumer credit to Personal Income	2
MZMSL	MZM Money Stock	6
DTCOLNVHFNM	Consumer Motor Vehicle Loans Outstanding	6
DTCTHFNM	Total Consumer Loans and Leases Outstanding	6
INVEST	Securities in Bank Credit at All Commercial Banks	6
FEDFUNDS	Effective Federal Funds Rate	2
CP3Mx	3-Month AA Financial Commercial Paper Rate	2
TB3MS	3-Month Treasury Bill:	2
TB6MS	6-Month Treasury Bill:	2
GS1	1-Year Treasury Rate	2
GS5	5-Year Treasury Rate	2
GS10	10-Year Treasury Rate	2
AAA	Moody's Seasoned Aaa Corporate Bond Yield	2
BAA	Moody's Seasoned Baa Corporate Bond Yield	2
COMPAPFFx	3-Month Commercial Paper Minus FEDFUNDS	1
TB3SMFFM	3-Month Treasury C Minus FEDFUNDS	1
TB6SMFFM	6-Month Treasury C Minus FEDFUNDS	1
T1YFFM	1-Year Treasury C Minus FEDFUNDS	1
T5YFFM	5-Year Treasury C Minus FEDFUNDS	1
T10YFFM	10-Year Treasury C Minus FEDFUNDS	1
AAAFFM	Moody's Aaa Corporate Bond Minus FEDFUNDS	1
BAAFFM	Moody's Baa Corporate Bond Minus FEDFUNDS	1
TWEXAFEGSMTHx	Trade Weighted U.S. Dollar Index	5

Table A.25: Macro Variables - Description cont'd

Notes: The table tabulates the variable acronym of the macro time-series, the description, as well as, the transformation code ("tcode").The transformation code denotes the variable transformation to make the series covariance stationary, (McCracken and Ng, 2016). The tcode corresponds to: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$.

Acronym	**Definition**	**tcode**
EXSZUSx	Switzerland / U.S. Foreign Exchange Rate	5
EXJPUSx	Japan / U.S. Foreign Exchange Rate	5
EXUSUKx	U.S. / U.K. Foreign Exchange Rate	5
EXCAUSx	Canada / U.S. Foreign Exchange Rate	5
WPSFD49207	PPI: Finished Goods	6
WPSFD49502	PPI: Finished Consumer Goods	6
WPSID61	PPI: Intermediate Materials	6
WPSID62	PPI: Crude Materials	6
OILPRICEx	Crude Oil, spliced WTI and Cushing	6
PPICMM	PPI: Metals and metal products:	6
CPIAUCSL	CPI : All Items	6
CPIAPPSL	CPI : Apparel	6
CPITRNSL	CPI : Transportation	6
CPIMEDSL	CPI : Medical Care	6
CUSR0000SAC	CPI : Commodities	6
CUSR0000SAD	CPI : Durables	6
CUSR0000SAS	CPI : Services	6
CPIULFSL	CPI : All Items Less Food	6
CUSR0000SA0L2	CPI : All items less shelter	6
CUSR0000SA0L5	CPI : All items less medical care	6
PCEPI	Personal Cons. Expend.: Chain Index	6
DDURRG3M086SBEA	Personal Cons. Exp: Durable goods	6
DNDGRG3M086SBEA	Personal Cons. Exp: Nondurable goods	6
DSERRG3M086SBEA	Personal Cons. Exp: Services	6
S&P 500	S&P's Common Stock Price Index: Composite	5
S&P: indust	S&P's Common Stock Price Index: Industrials	5
S&P div yield	S&P's Composite Common Stock: Dividend Yield	2
S&P PE ratio	S&P's Composite Common Stock: Price-Earnings Ratio	5
VXOCLSx	VXO	1

Table A.26: Macro Variables - Description cont'd

Notes: The table tabulates the variable acronym of the macro time-series, the description, as well as, the transformation code ("tcode").The transformation code denotes the variable transformation to make the series covariance stationary, (McCracken and Ng, 2016). The tcode corresponds to: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $log(x_t)$; (5) $\Delta log(x_t)$; (6) $\Delta^2 log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$.

A.7 Results for monthly return prediction using XGBoost

A typical application of xgboost on the whole feature set, including all macro features, and with GPU acceleration for xgboost, shows there is little predictability in monthly excess return data.

XGBoost	bond portfolios	indiv. bonds
Best val. OOS-R2	0.04	-0.01
Test OOS	**0.003**	**-0.006**
Start date testing	01/2004	01/2006
most important feature	med_s2p	med_lme
2nd important feature	med_r2	med_op
3rd important feature	avg_ereturn_treasury	GS1

Table A.27: Prediction via XGBoost - Results