The Corporate Bond Factor Zoo

Alexander Dickerson, Christian Julliard, and Philippe Mueller*

February 25, 2024

Abstract

Analyzing 563 trillion possible models, we find that the majority of tradable factors designed to price bond markets are unlikely sources of priced risk, and only one novel tradable bond factor, capturing the bond post-earnings announcement drift, should be included in the stochastic discount factor (SDF) with very high probability. Nevertheless, the SDF is dense in the space of observable factors, with both nontradable and equity-based ones being salient for pricing corporate bonds. A Bayesian model averaging–SDF explains corporate risk premia better than all existing models, both in- and out-of-sample, and captures business cycle and market crash risks.

Keywords: Corporate bonds; Factor zoo; Factor models; Asset pricing; Bayesian methods.
JEL classification: G10; G12; G40; C12; C13; C52.

*Alexander Dickerson, School of Banking & Finance, The University of New South Wales, alexander.dickerson1@unsw.edu.au; Christian Julliard, Department of Finance, London School of Economics, c.julliard@lse.ac.uk; and Philippe Mueller (corresponding author), Warwick Business School, The University of Warwick, Scarman Rd, Coventry CV4 7AL, UK, philippe.mueller@wbs.ac.uk. We gratefully acknowledge comments and suggestions from Svetlana Bryzgalova, Mikhail Chernov, Jiantao Huang, Ian Martin, and Mamdouh Medhat. We gratefully acknowledge financial support from INQUIRE Europe. The companion website to this paper, openbondassetpricing.com, contains updated corporate bond factor data.

Wherever there is risk, it must be compensated to the lender by a higher premium or interest.

— J. R. McCullough (1830, pp. 508–9)

Despite its size and first order relevance for firm financing, after six decades of empirical research, there is still very little agreement about the sources of risk that drive prices in the corporate bond market. Leveraging recent advances in Bayesian econometrics, we fill this gap and provide a comprehensive analysis of all the factors and models proposed to date, as well as their possible combinations, and their interplay with asset pricing factors identified in the equity market literature. This allows us to pinpoint both the robust sources of priced risk, and a novel benchmark stochastic discount factor (SDF) that prices corporate bonds, in- and out-of-sample, significantly better than all existing models.

The early literature documents risk factors, meant to capture economy-wide conditions, attitude toward risk, and firm characteristics, common to stocks and bonds. However, more recent studies propose alternative factors, frequently derived from interest rate term structures and specific bond metrics, which diverge from those typically associated with the stock market. That is, the corporate bond space develops its own "factor zoo," independent from the equity zoo and its associated models, as the latter are considered insufficient to explain corporate bond risk premia. More recently, the accepted wisdom is again called into question. First, what had arguably emerged as the benchmark model for bonds (Bai, Bali, and Wen (2019)), is now retracted due to lead and lag errors present in the factors. Second, using misspecification-robust inference, Dickerson, Mueller, and Robotti (2023) show that the low dimensional factor models in the literature add little spanning to a simple *bond* version of the Capital Asset Pricing Model (the CAPMB). Third, van Binsbergen, Nozawa, and Schwert (2023) find that, once returns are adjusted for duration risk, the simple (equity) CAPM has higher explanatory power for corporate bonds than benchmark models. That is, the overall understanding of risks priced in the corporate bond market has not progressed much beyond the theoretical milestones of decades past.

We analyze empirically over 562 trillion models stemming from the joint zoo of corporate bond and equity factors, and we do so while relaxing the cornerstone assumptions of previous studies: the existence of a unique, low-dimensional, correctly specified and well identified factor model. First, we find that the "true" latent SDF of corporate bonds is *dense* in the space of observable bond and equity factors—literally dozens of factors, both tradable and nontradable, are necessary to span the risks driving bond prices. Yet, the SDF-implied maximum Sharpe ratio is not excessive, indicating that multiple factors load on common sources of fundamental risk. Importantly, density of the SDF implies that the sparse models considered in the previous literature are affected by severe misspecification and, as we show, outperformed by the most likely (four to eight) SDF components that we identify.

Second, a Bayesian model averaging–stochastic discount factor (BMA-SDF) over the space of all possible models (including bond, equity, and nontradable factors) explains corporate bond risk premia better than all existing models and most likely factors, both in- and out-of-sample. Moreover, the BMA-SDF has a clear business cycle pattern: it increases during expansions and peaks right before recessions and around the time of financial market crashes. That is, the estimated SDF behaves as one would expect from the intertemporal marginal rate of substitution of an agent exposed to the risks arising from general economic conditions and market turmoil.

Third, we show that the majority of tradable factors designed to price corporate bonds are unlikely sources of priced risk. However, we find compelling evidence that a novel (tradable) factor, theoretically motivated yet never used before for asset pricing, and capturing the bond post-earnings announcement drift (PEADB), should be included in the SDF with very high probability.[1] Furthermore, two nontradable factors, meant to capture inflation risk (Kang and Pflueger (2015)) and the slope of the term structure of interest rates (Koijen, Lustig, and Van Nieuwerburgh (2017)), a well documented predictor of economic activity, are likely components of the SDF. Similarly, a broad based corporate bond market index is also likely to be part of the SDF (albeit the single factor CAPMB is rejected by the data). Moreover, when expanding the set of candidate pricing factors to include equity-based ones, measures of firm size (Fama and French (1992), Daniel, Mota, Rottke, and Santos (2020)), market liquidity (Pástor and Stambaugh (2003)), and long term reversal (Jegadeesh and

[1] The post-earnings announcement drift phenomenon is the observation, first documented in equity markets, that firms that experience positive earnings surprises subsequently earn higher returns than those with negative earnings surprises. See, e.g., Hirshleifer and Teoh (2003), Della Vigna and Pollet (2009), Hirshleifer, Lim, and Teoh (2011) and Nozawa, Qiu, and Xiong (2023) for the microfoundations of this phenomenon.

Titman (2001)) also have posterior probabilities of being part of the SDF that exceed (albeit in some cases only marginally) their prior values.

Fourth, beside the individual factors mentioned above, both nontradable and equity-based factors *jointly* play an important role in the BMA-SDF, and are more likely components of the pricing measure than all other bond-based tradable factors. Nevertheless, several factors are only weakly identified in the cross-section of bond returns, and, hence, invalidate canonical inference (see, e.g., Gospodinov, Kan, and Robotti (2014)). This, importantly, is *not* a problem for the Bayesian method that we employ, since it is by design robust to weak factors (Bryzgalova, Huang, and Julliard (2023)) and, as we show, successfully shrinks the market price of risk of weak and (likely) spurious factors toward zero.

Remarkably, all of the above results (SDF factor density, BMA-SDF pricing ability, strong evidence in favour of PEADB, identity of most likely factors and their types) are extremely stable across data sources and sample periods, and independently of whether test asset returns are computed in excess of the short term risk free rate or a duration-matched U.S. Treasury Bond rate. Nevertheless, the outperformance of the BMA-SDF compared to existing models, and the salient role of equity factors for pricing corporate bonds in cross-sectional out-of-sample exercises, is even more evident when considering duration-adjusted bond returns. Furthermore, the out-of-sample pricing performance of the BMA-SDF is very stable across 127 different cross-sections of test assets (virtually spanning the entire universe considered in the previous literature).

Globally, the total market capitalizations of bond and equity markets are almost on par at about USD 125 trillion each. The total size of the corporate bond market is around USD 40 trillion, with the United States accounting for 27% thereof. That is, in market value terms, corporate bonds are about a third of the size of equities. However, from an asset manager's perspective, it is arguably equally import to understand what drives bond and stock prices, as corporate bonds are usually held by institutional investors or via managed investment vehicles (see, e.g., Boyarchenko, Elias, and Mueller 2023). Understanding what factors, and models, are relevant for pricing the universe of corporate bonds in the U.S. (and the rest of the world) is salient for investors, as it guides how investment portfolios of these assets should be formed in the first place. Importantly, our results can be used directly to motivate and implement smart beta strategies for corporate bond portfolios.

Furthermore, the complete market benchmark states that the pricing measure should be consistent across asset classes, hence a systematic exploration of the sources of risk priced in corporate bonds should encompass not only factors proposed in the bond market literature, but also those suggested for the equity space—that is, one should parse the *joint* factor zoo of bonds and equities, and that is exactly what we deliver.

Note that, unlike equities, where factor data is mostly readily available, a reliable public database of corporate bond factors does not exist. To address this pitfall for conducting rigorous and replicable asset pricing in bonds, we use best in class data and make all factors and code available on openbondassetpricing.com. Moreover, to avoid the litany of potential data errors arising from the Trade Reporting and Compliance Engine (TRACE) transaction-based data, we use the industry-grade data from the Intercontinental Exchange (ICE) and the Lehman Brothers Database over the combined 1986:01 to 2021:09 sample.

Closely related literature. Our research contributes to the active and growing body of work that critically reevaluates existing findings in the empirical asset pricing literature using robust inference methods. Following Harvey, Liu, and Zhu (2016), a large body of literature has tried to understand which existing factors (or their combinations) drive the cross-section of (equity) returns. In particular, Gospodinov, Kan, and Robotti (2014) develop a general method for misspecification-robust inference, while Giglio and Xiu (2021) exploit the invariance principle of the PCA and recover the price of risk of a given factor from the projection on the span of latent factors driving a cross-section of returns. Similarly, Dello-Preite, Uppal, Zaffaroni, and Zviadadze (2023) recover latent factors from the residuals of an asset pricing model, effectively completing the span of the SDF. Feng, Giglio, and Xiu (2020) combine cross-sectional asset pricing regressions with the double-selection LASSO of Belloni, Chernozhukov, and Hansen (2014) to provide valid inference on the selected sources of risk when the true SDF is sparse. Kozak, Nagel, and Santosh (2020) uses a ridge-based approach to approximate the SDF and compare sparse models based on principal components of returns with sparse models based on characteristics. Our approach instead identifies a dominant pricing model—if such model exists—or a BMA across the space of all models, even if the true model is not sparse in nature, hence cannot be proxied by a small number of factors.

As Harvey (2017) stresses in his American Finance Association presidential address, the factor zoo naturally calls for a Bayesian solution—and we adopt one. In particular, we leverage the Bayesian method for model estimation, selection, and averaging developed in Bryzgalova, Huang, and Julliard (2023). Numerous strands of the literature rely on Bayesian tools for asset allocation, model selection, and performance evaluation. Our approach is most closely linked to Pástor and Stambaugh (2000) and Pástor (2000) in that we assign a prior distribution to the vector of pricing errors, and this maps into a natural and transparent prior for the maximal Sharpe ratio achievable in the economy. Barillas and Shanken (2018) also extend the prior formulation of Pástor and Stambaugh (2000) and provide a closed-form solution for the Bayes factors when all factors are tradable in nature. Chib, Zeng, and Zhao (2020) show that the improper prior formulation of Barillas and Shanken (2018) is problematic, and provide a new class of priors that leads to valid comparison for traded factor models. As in these papers, our model and factor selection is based on posterior probabilities, but our method is designed to work with both tradable and *nontradable* factors. Most importantly, our approach can deal with a very large factor space, is not affected by the common identification failures that invalidate inference in asset pricing, and provides an optimal method for aggregating the pricing information stemming from all factors.[2]

As widely documented, ratings and default risk proxies are insufficient to explain corporate bond risk premia (see, e.g., Elton, Gruber, Agrawal, and Mann (2004) and Driessen (2005)). Furthermore, in the complete market benchmark the pricing measure should be consistent across asset classes, and equilibrium models normally yield nontradable state variables. Hence, we consider a very broad collection of potential sources of risk that goes well beyond the set of bond tradable factors that have been the main focus of a large part

[2]BMA is an optimal aggregation procedure for a very wide set of optimality criteria (see, e.g., Raftery and Zheng (2003) and Schervish (1995)). In particular, it is "optimal on average," i.e., no alternative method can outperform the BMA for all values of the true unknown parameters. Avramov, Cheng, Metzker, and Voigt (2023) also propose a framework to integrate factor models via posterior probabilities in the presence of model uncertainty, but their approach is only appropriate for tradable factors and is not designed to be robust to the identification and inference problems arising from weak factors—problems that, as shown in Bryzgalova, Huang, and Julliard (2023), cannot be solved by simply projecting nontradable factors on the space of returns and then performing inference using the resulting mimicking portfolios.

of the corporate bond literature (see Dickerson, Mueller, and Robotti (2023), Dickerson, Robotti, and Rossetti (2023) and Dick-Nielsen, Feldhütter, Pedersen, and Stolborg (2023) for critical reassessments of these factors).

In particular, our paper is related to the body of work that explores whether equity market risk proxies (see, e.g., Blume and Keim (1987), Fama and French (1993) and Elton, Gruber, Agrawal, and Mann (2001)), equity volatilities (see, e.g., Campbell and Taksler (2003) and Chung, Wang, and Wu (2019)), and equity-based characteristics (see, e.g., Fisher (1959), Giesecke, Longstaff, Schaefer, and Strebulaev (2011)) are likely drivers of corporate bond returns. Overall, we find that not only several equity-based factors have posterior probabilities of being part of the SDF that exceed their prior values, but also that equity factors jointly carry relevant information for pricing corporate bonds.

Several theoretical contributions stress that real economic activity and the business cycle should be among the drivers of bond risk premia (see, e.g., Bhamra, Kuehn, and Strebulaev (2010), Khan and Thomas (2013), Chen, Cui, He, and Milbradt (2018), and Favilukis, Lin, and Zhao (2020)). Echoing both the general equilibrium model predictions of Gomes and Schmid (2021) and the empirical findings of Elton, Gruber, and Blake (1995) and Elkamhi, Jo, and Nozawa (2023), we show that the BMA-SDF has a clear business cycle pattern and peaks right before recessions and at times of market crashes, and that nontradable factors (especially proxies of the economic cycle such as the slope of the yield curve), are salient components of the pricing measure.[3] Furthermore, in line with both the model predictions and international empirical evidence of Kang and Pflueger (2015) (see also Bhamra, Fisher, and Kuehn (2011) and Ceballos (2022)), we find that inflation volatility risk is a likely component of the SDF that drives corporate bond prices.

Our work also relates to behavioural biases and market frictions in asset pricing. In particular, mirroring the evidence of Daniel, Hirshleifer, and Sun (2020) and Bryzgalova, Huang, and Julliard (2023) in the equity market, we show that the post earning announce-

[3]Elton, Gruber, and Blake (1995) show that adding fundamental macro-risk variables (such as GNP, inflation and term spread measures) significantly improves pricing performance relative to equity and bond market index models. Elkamhi, Jo, and Nozawa (2023) show that the long-run consumption risk measure of Parker and Julliard (2003) yields a one-factor model with significant explanatory power for corporate bonds, and such an SDF, as documented in Parker and Julliard (2005), has a very strong business cycle pattern.

ment drift of bonds (see Nozawa, Qiu, and Xiong (2023)) is an extremely likely driver of corporate bond risk premia. Furthermore, we find empirical support, albeit much weaker, for the (equity) long-term reversal factor of Jegadeesh and Titman (2001). And finally, we find some evidence for market liquidity carrying explanatory power for the cross-section of bond returns (see, e.g., Bao, Pan, and Wang (2011) and Lin, Wang, and Wu (2011)), in that the inclusion of a nontradable liquidity proxy (Pástor and Stambaugh (2003)) in the SDF is, at least marginally, supported by the data.

1 Data

We rely on the constituents of the corporate bond data set from the Bank of America Merrill Lynch (BAML) High Yield (H0A0) and Investment Grade (C0A0) indices made available via the Intercontinental Exchange (ICE), which starts in January 1997 and ends in September 2021. For the period from January 1986 to December 1996 we use the Lehman Brothers Fixed Income Database (LHM). This data is then merged with the Mergent Fixed Income Securities Database (FISD), which contains additional bond characteristics. We follow van Binsbergen, Nozawa, and Schwert (2023) and begin the LHM sample in 1986, the first year with at least 100 high-yield bonds per month in the sample.[4] After merging the two data sets and applying the standard filters (discussed below), our bond-level data spans almost 36 years over the period 1986:01 to 2021:09 ($T = 429$ months) and comprises over 30,000 unique bonds. A detailed description of the databases and associated cleaning procedures is available in Section IA.1 of the Internet Appendix

In contrast to the Trade Reporting and Compliance Engine (TRACE) transaction-based data, ICE data have three key advantages. First, monthly prices in the ICE data are sampled exactly at the end of each month, which means monthly returns always use month-end prices. This is particularly important when using equity-characteristics as inputs to compute bond factor returns. Aligning the timing avoids potential look-ahead bias or lead-lag errors and

[4] The start date of January 1986 is broadly consistent with prior literature that utilises the LHM data. See Eom, Helwege, and Huang (2004), Feldhütter and Schaefer (2018), and Avramov, Chordia, Jostova, and Philipov (2022). Prior to 1986, bonds in LHM are predominantly investment grade (91% of bonds) with 67% of all bonds priced with matrix pricing (i.e., the prices are not actual dealer quotes).

ensures that both stock and bond returns are computed with prices that are sampled on the last day of each month. In contrast, this is not the case for the TRACE data, where a bond must trade (i.e., transact) sufficiently close to the end of the month to be included in the monthly data set. Given that on average bonds do not trade on roughly 68% of days (see, e.g., Palhares and Richardson (2020)), computing monthly returns with daily TRACE data is complicated by using prices that are not sampled at month end or are missing altogether. A further consequence of this feature of the data is a larger sample size of ICE relative to the Wharton Research Data Services (WRDS) version of TRACE.

Second, the effects of market microstructure noise are, for the most part, removed, as ICE only provides the bid-side of the trade (or quote if no trade occurs).[5] Furthermore, the effects of potential data errors and ad hoc data decisions are inherently removed because the data are provided (pre-processed) by an institutional grade data provider.

Finally, ICE provides pre-computed corporate bond characteristics including the total bond return, duration-adjusted return, bond duration and convexity, as well as several proxies for credit spreads taking into account bond optionality and different risk-free rate benchmarks. This makes ICE particularly convenient for empirical analysis and alleviates any potential error in computing these metrics, i.e., it introduces homogeneity into the results such that they can be trusted.

Because of the above considerations, and as noted by Kelly, Palhares, and Pruitt (2023), ICE has become the de facto 'gold standard' for corporate bond empirical studies and is (or should be) the primary data source for the corporate bond literature.[6]

We apply the following standard filters to the bond data: i) We remove bonds that are not publicly traded in the U.S. market. These include bonds issued through private placement, bonds issued under Rule 144A, bonds that are not traded in USD, and bonds from issuers not based in the U.S. ii) We remove bonds that are classified as structured notes, mortgage backed or asset backed, agency backed, equity linked or convertible. iii) We exclude bonds

[5]Although the level of market microstructure noise (MMN) has consistently declined since the introduction of the TRACE system, Dickerson, Robotti, and Rossetti (2023) show that it can still adversely affect the measurement of corporate bond price-based anomaly characteristics.

[6]See Kelly and Pruitt (2022) and Andreani, Palhares, and Richardson (2023) for detailed discussions on the differences and similarities between ICE and TRACE data.

that have a floating coupon rate. iv) Finally, we exclude bonds that have less than one year remaining until maturity. For robustness, we replicate our main results using the TRACE data processed by the WRDS data science team over a shorter sample beginning in 2002:09, and show that they remain materially unchanged.

Corporate bond returns. In the baseline analysis, we specify *excess* bond returns as the total bond return minus the one-month risk-free rate of return.[7] In addition, we follow the advice of van Binsbergen, Nozawa, and Schwert (2023) and repeat our analysis with 'duration-adjusted' returns, where the bond *excess* return is computed as the total bond return minus a portfolio of duration-matched U.S. Treasury Bond returns. Details of the duration adjustment are provided in Appendix C. Note that we do not further winsorize, trim, or tamper with the underlying bond return data in any way, avoiding the biases that such procedures normally induce.

Corporate bond factors and anomalies. Our bond-specific factor zoo includes 14 traded bond factors and 11 nontraded factors.[8] From the equity literature, we include an additional 24 traded factors. Overall, we consider a total of 49 factors, of which 38 are traded and 11 are nontraded. We provide an overview of the factors in Table A.1 of Appendix B. All of the traded equity and nontraded factors are publicly available from the various authors' personal websites listed therein.[9]

Corporate bond test asset portfolios. For our cross-sectional analyses at the portfolio-level, we construct a set of bond portfolios that are sorted on various bond characteristics. To ensure a broad enough cross-section for our in-sample estimation of the BMA, we use 50 bond portfolios. The first 25 portfolios are double-sorted on credit spreads and bond size. The remaining 25 portfolios are double-sorted on bond ratings and time-to-maturity. All portfolios are value-weighted by the amount outstanding of the bond issue, defined as the dollar par value multiplied by the number of outstanding units of the bond.

[7] We source the one-month risk-free rate from Kenneth French's website.

[8] Many of the nontraded factors employed for bonds have also been used for stocks.

[9] We make our 14 traded bond factors available on the companion website: openbondassetpricing.com

The chosen characteristics yield a significant dispersion of average in-sample bond portfolio returns. The inclusion of portfolios sorted on credit spreads is motivated by the work of Nozawa (2017) who finds that bond credit spreads are an important driver of the cross-sectional variation in excess corporate bond returns.[10] Bond ratings are provided by Standard & Poors (S&P) and are a fundamental characteristic of bonds. They underpin most traded bond factors, define institutional investment guidelines, and capture default risk.

Finally, we include the traded bond factors as additional test assets since, as emphasized in Barillas and Shanken (2016), factors included in a model should price any factors excluded from the model. This, along with the use of the nonspherical pricing error formulation (i.e., GLS) also imposes (asymptotically) the restriction of factors pricing themselves. Overall, the cross-section contains a broad array of 64 portfolios, sorted on well-known bond characteristics and the underlying traded bond factors themselves.

Out-of-sample bond portfolios To test the asset pricing efficacy of the BMA-SDF estimated on the in-sample test assets, we specifically construct a broad cross-section of additional bond portfolios using bond characteristics that are *not* used to construct the in-sample portfolios. To do so, we include decile-sorted portfolios on bond historical 95% value-at-risk, duration, bond value (Houweling and Van Zundert (2017)), bond book-to-market (Bartram, Grinblatt, and Nozawa (2020)), long-term reversals (Bali, Subrahmanyam, and Wen (2021)), momentum (Jostova, Nikolova, Philipov, and Stahel (2013)), as well as the 17 Fama French industry portfolios (following Lewellen, Nagel, and Shanken (2010)), for a total of 77 portfolios.

2 Econometric method

This section introduces the notation and summarises the methods employed in in our empirical analysis. We consider linear factor models for the Stochastic Discount Factor (SDF). We focus on the SDF representation since we aim to identify the factors that have

[10]We follow the credit spread portfolio formation method in Elkamhi, Jo, and Nozawa (2023) and construct the portfolios based on the average bond credit spreads between months $t-12$ and $t-1$.

pricing ability for the cross-section of corporate bonds returns.[11]

The returns of N test assets, which are long-short portfolios, are denoted by $\boldsymbol{R_t} = (R_{1t} \dots R_{Nt})^\top$, $t = 1, \dots T$. We consider K factors, $\boldsymbol{f}_t = (f_{1t} \dots f_{Kt})^\top$, $t = 1, \dots T$, that can be either tradable or nontradable. A linear SDF takes the form $M_t = 1 - (\boldsymbol{f}_t - \mathbb{E}[\boldsymbol{f}_t])^\top \boldsymbol{\lambda_f}$, where $\boldsymbol{\lambda_f} \in \mathbb{R}^K$ is the vector containing the market prices of risk associated with the individual factors. Throughout the paper, $\mathbb{E}[X]$ or μ_X denote the unconditional expectation of arbitrary random variable X.

In the absence of arbitrage opportunities, we have that $\mathbb{E}[M_t \boldsymbol{R_t}] = \boldsymbol{0_N}$, hence expected returns are given by $\boldsymbol{\mu_R} \equiv \mathbb{E}[\boldsymbol{R}_t] = \boldsymbol{C_f \lambda_f}$, where $\boldsymbol{C_f}$ is the covariance matrix between $\boldsymbol{R}_t$ and $\boldsymbol{f}_t$, and prices of risk ($\boldsymbol{\lambda_f}$) are commonly estimated via the cross-sectional regression

$$\boldsymbol{\mu_R} = \lambda_c \boldsymbol{1_N} + \boldsymbol{C_f \lambda_f} + \boldsymbol{\alpha} = \boldsymbol{C\lambda} + \boldsymbol{\alpha}, \tag{1}$$

where $\boldsymbol{C} = (\boldsymbol{1_N}, \boldsymbol{C_f})$, $\boldsymbol{\lambda}^\top = (\lambda_c, \boldsymbol{\lambda_f}^\top)$, λ_c is a scalar average mispricing (equal to zero under the null of the model being correctly specified), $\boldsymbol{1_N}$ is an N-dimensional vector of ones, and $\boldsymbol{\alpha} \in \mathbb{R}^N$ is the vector of pricing errors in excess of λ_c (also equal to zero under the null of the model).

Such models are usually estimated via GMM, MLE or two-pass regression methods (see, e.g., Hansen 1982, Cochrane 2005). Nevertheless, as pointed out in a large literature, the underlying assumptions for the validity of these methods (see, e.g., Newey and McFadden 1994), are often violated (see, e.g., see Kleibergen and Zhan 2020 and Gospodinov and Robotti 2021), and identification problems arise in the presence of a *weak* factor (i.e., a factor that does not have enough comovement with any of the assets, or has very little cross-sectional dispersion in this comovement, but is nonetheless considered a part of the SDF). These issues in turn lead to wrong inference for both weak and strong factors, erroneous model selection, and inflate the canonical measures of model fit.[12]

Albeit robust frequentist inference methods have been suggested in the literature for

[11] Recall that a factor might have a significant risk premium even if it is not part of the SDF, just because it has non-zero correlation with the true latent SDF. Hence, in order to identify the pricing measure, focusing on the SDF representation is the natural choice.

[12] These problems are common to GMM (Kan and Zhang, 1999a), MLE (Gospodinov, Kan, and Robotti, 2019), Fama-MacBeth regressions (Kan and Zhang 1999b, Kleibergen 2009), and even Bayesian approaches with flat priors for risk prices (Bryzgalova, Huang, and Julliard, 2023).

specific settings, our task is complicated by the fact that we want to parse the entire zoo of bond factors, rather than estimate and test an individual model. Furthermore, we aim to identify the best specification—*if* a dominant model exist—or aggregate the information in the factor zoo into a single SDF if no clear best model arises. Therefore, we rely on the Bayesian method proposed recently in Bryzgalova, Huang, and Julliard (2023), since it is applicable to both tradable and nontradable factors, can handle the entire factor zoo, is valid under misspecification, and is robust to weak inference problems. This Bayesian approach is conceptually simple, since it leverages the naturally hierarchical structure of cross-sectional asset pricing, and restores the validity of inference using transparent and economically motivated priors.

Consider first the time-series layer of the estimation problem. Without loss of generality, we order the K_1 tradable factors first ($\boldsymbol{f}_t^{(1)}$), followed by K_2 nontradable factors ($\boldsymbol{f}_t^{(2)}$), hence $\boldsymbol{f}_t \equiv (\boldsymbol{f}_t^{(1),\top}, \boldsymbol{f}_t^{(2),\top})^\top$ and $K_1 + K_2 = K$. Denote with $\boldsymbol{Y_t} \equiv \boldsymbol{f_t} \cup \boldsymbol{R_t}$ the union of factors and returns, where $\boldsymbol{Y_t}$ is a p-dimensional vector.[13] Modelling $\{\boldsymbol{Y_t}\}_{t=1}^T$ as multivariate Gaussian with mean $\boldsymbol{\mu_Y}$ and variance matrix $\boldsymbol{\Sigma_Y}$, and adopting the conventional diffuse prior $\pi(\boldsymbol{\mu_Y}, \boldsymbol{\Sigma_Y}) \propto |\boldsymbol{\Sigma_Y}|^{-\frac{p+1}{2}}$, yields the canonical Normal-inverse-Wishart posterior for the time series parameters $(\boldsymbol{\mu_Y}, \boldsymbol{\Sigma_Y})$ in equations (A.7)-(A.8) of Appendix A.

The cross-sectional layer of the inference problem allows for misspecification of the factor model via the average pricing errors $\boldsymbol{\alpha}$ in equation (1). We model these pricing errors, as in the previous literature (e.g., Pástor and Stambaugh 2000, Pástor 2000), as $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{0_N}, \sigma^2 \boldsymbol{\Sigma_R})$, yielding the cross-sectional likelihood (conditional on the time series parameters)

$$p(\text{data}|\boldsymbol{\lambda}, \sigma^2) = (2\pi\sigma^2)^{-\frac{N}{2}} |\boldsymbol{\Sigma_R}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2} (\boldsymbol{\mu_R} - \boldsymbol{C\lambda})^\top \boldsymbol{\Sigma_R}^{-1} (\boldsymbol{\mu_R} - \boldsymbol{C\lambda}) \right\}, \tag{2}$$

where in the cross-sectional regression the "data" are the expected risk premia, $\boldsymbol{\mu_R}$, and the factor loadings, $\boldsymbol{C} \equiv (\boldsymbol{1_N}, \boldsymbol{C_f})$. The above likelihood can then be combined with a prior for risk prices (presented below) to obtain a posterior distribution and guide inference and model selection.

To handle model and factor selection we introduce a vector of binary latent variables

[13] If one requires the tradable factors to price themselves, then $\boldsymbol{Y_t} \equiv (\boldsymbol{R_t}^\top, \boldsymbol{f}_t^{(2),\top})^\top$ and $p = N + K_2$.

$\boldsymbol{\gamma}^\top = (\gamma_0, \gamma_1, \ldots, \gamma_K)$, where $\gamma_j \in \{0, 1\}$. When $\gamma_j = 1$, the j-th factor (with associated loadings $\boldsymbol{C_j}$) should be included in the SDF, and should be excluded otherwise.[14] In the presence of potentially weak factors, and hence unidentified prices of risk, the posterior probabilities of models and factors are not well defined under flat priors. Hence, we introduce a (economically motivated) prior that, albeit not informative, restores the validity of posterior inference (see Bryzgalova, Huang, and Julliard 2023). In particular, we model the uncertainty underlying the estimation and model selection problem with a (continuous spike-and-slab) mixture prior, $\pi(\boldsymbol{\lambda}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega}) = \pi(\boldsymbol{\lambda} \mid \sigma^2, \boldsymbol{\gamma})\pi(\sigma^2)\pi(\boldsymbol{\gamma} \mid \boldsymbol{\omega})\pi(\boldsymbol{\omega})$, where

$$\lambda_j \mid \gamma_j, \sigma^2 \sim \mathcal{N}\left(0, r(\gamma_j)\psi_j\sigma^2\right). \tag{3}$$

Note the presence of three new elements, ψ_j, $r(\gamma_j)$, and $\pi(\boldsymbol{\omega})$, in the prior formulation.[15]

First, $r(\gamma_j)$ captures the "spike-and-slab" nature of the prior formulation. When the factor should be included, we have $r(\gamma_j = 1) = 1$, and the prior, the "slab", is just a diffuse distribution centred at zero. When instead the factor should not be in the model, $r(\gamma_j = 0) = r \ll 1$, the prior is extremely concentrated—a "spike" at zero. As $r \to 0$, the prior spike is just a Dirac distribution at zero, hence it removes the factor from the SDF.[16]

Second, ψ_j is a function of the data that (endogenously) penalises, like a shrinkage estimator, factors that are likely to be causing identification failure:

$$\psi_j = \psi \times \widetilde{\boldsymbol{\rho}}_j^\top \widetilde{\boldsymbol{\rho}}_j, \tag{4}$$

where $\widetilde{\boldsymbol{\rho}}_j \equiv \boldsymbol{\rho}_j - \left(\frac{1}{N}\sum_{i=1}^N \rho_{j,i}\right) \times \boldsymbol{1_N}$, $\boldsymbol{\rho_j}$ is an $N \times 1$ vector of correlation coefficients between factor j and the test assets, and $\psi \in \mathbb{R}_+$ is a tuning parameter that controls the degree of shrinkage across all factors. That is, factors that have vanishing correlation with asset returns, or extremely low cross-sectional dispersion in their correlations (hence cannot help in explaining cross-sectional differences in returns), have a low value of ψ_j and are therefore endogenously shrunk toward zero. Instead, such prior has no effect on the estimation of

[14] Note that we always include the common intercept in the cross-sectional layer, that is, $\gamma_0 = 1$ always.

[15] For the cross-sectional variance scale parameter σ^2 we assume the customary diffuse prior $\pi(\sigma^2) \propto \sigma^{-2}$. As per Proposition 1 of Chib, Zeng, and Zhao (2020), since the parameter σ is common across models and has the same support in each model, the marginal likelihoods obtained under this improper prior are valid and comparable.

[16] We set $r = 0.001$ in our empirical analysis.

strong factors since these have large and disperse correlations with the test assets, yielding a large ψ_j and consequently a diffuse prior.

Third, the prior $\pi(\boldsymbol{\omega})$ not only gives us a way to sample from the space of potential models, but also encodes belief about the sparsity of the true model using the prior distribution $\pi(\gamma_j = 1|\omega_j) = \omega_j$. Following the literature on predictors selection, we set

$$\pi(\gamma_j = 1|\omega_j) = \omega_j, \quad \omega_j \sim Beta\left(a_\omega, b_\omega\right). \tag{5}$$

Different hyperparameters a_ω and b_ω determine whether one a priori favors more parsimonious models or not. The prior expected probability of selecting a factor is simply $\frac{a_\omega}{a_\omega+b_\omega}$. We set $a_\omega = b_\omega = 1$ in the benchmark case, that is, we have a uniform (hence flat) prior for the model dimensionality and each factor has an ex ante expected probability of being selected equal to 50%.[17]

Furthermore, note that the only free "tuning" parameter in our setting, ψ in equation (4), has a straightforward economic interpretation, since the expected prior Sharpe ratio (SR) achievable with the factors is just $\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2] = \frac{1}{2}\psi\sigma^2 \sum_{k=1}^{K} \tilde{\boldsymbol{\rho}}_k^\top \tilde{\boldsymbol{\rho}}_k$ as $r \to 0$.[18] That is, in our empirical analysis we report results for various prior expectations of the Sharpe ratio achievable in the economy.[19]

The above hierarchical system yields a well defined posterior distribution from which all the unknown parameters and quantities of interest (e.g., R^2, SDF-implied Sharpe ratio, and model dimensionality), can be sampled to compute posterior means and credible intervals via the Gibbs sampling algorithm in Appendix A. Most importantly, these posterior draws can be used to compute posterior model and factor probabilities, and hence identify robust sources of priced risk and—*if* such model exists—a dominant model for pricing corporate bonds.

Model and factor probabilities can also be used for aggregating optimally, rather then selecting, the pricing information in the factor zoo. For each possible model $\boldsymbol{\gamma}^m$ that one

[17]However, we could set for instance, $a_\omega = 1$ and $b_\omega >> 1$ to favor sparser models.

[18]Without a uniform prior for the SDF dimensionality the prior Sharpe ratio value becomes $\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2] = \frac{a_\omega}{a_\omega+b_\omega}\psi\sigma^2 \sum_{k=1}^{K} \tilde{\boldsymbol{\rho}}_k^\top \tilde{\boldsymbol{\rho}}_k$ as $r \to 0$. Hence, beliefs about the prior Sharpe ratio and model dimensionality fully pin down our hyperparameters.

[19]More precisely, we report results for different prior values of $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$.

could construct with the universe of factors, we have the corresponding SDF: $M_{t,\boldsymbol{\gamma}^m} = 1 - (\boldsymbol{f}_{t,\boldsymbol{\gamma}^m} - \mathbb{E}[\boldsymbol{f}_{t,\boldsymbol{\gamma}^m}])^\top \boldsymbol{\lambda}_{\boldsymbol{\gamma}^m}$. Therefore, we construct a Bayesian Model Averaging SDF (BMA-SDF) by averaging all possible SDFs using as weights the posterior probability of each model:[20]

$$M_t^{BMA} = \sum_{m=1}^{\bar{m}} M_{t,\boldsymbol{\gamma}^m} \Pr\left(\boldsymbol{\gamma}^m|\text{data}\right), \tag{6}$$

where $\bar{m}$ is the total number of possible models.

The BMA aggregates information about the true latent SDF over the space of all possible models, rather than conditioning on a particular model. At the same time, if a dominant model exists (a model for which $\Pr\left(\boldsymbol{\gamma}^m|\text{data}\right) \approx 1$), the BMA will use that model alone. Moreover, BMA aggregation is optimal under a wide range of criteria, but in particular, it is "optimal on average:" no alternative estimator can outperform it for all possible values of the true unknown parameters.[21] Furthermore, since its predictive distribution minimizes the Kullback-Leibler information divergence relative to the true unknown data-generating process, the BMA aggregation delivers the most likely SDF given the data, and the estimated density is as close as possible to the true unknown one, even if all of the models considered are misspecified.

3 Bayesian analysis of linear SDFs

In this section, we apply the hierarchical Bayesian method to a large set of factors proposed in the previous bond and equity literature. We start by focusing on factors specifically proposed to price the cross-section of corporate bonds in Section 3.1. In Section 3.2 we then proceed to further include factors that successfully price the cross-section of equity returns and highlight the factors that also contain relevant information for corporate bonds. Overall, we consider 38 tradable and 11 nontradable factors, yielding $2^{49} \approx 563$ trillion possible models for the combined bond and stock factor zoo, and $2^{25} \approx 33.6$ million models for the bond-based factors only.

[20]See, e.g., Raftery, Madigan, and Hoeting (1997) and Hoeting, Madigan, Raftery, and Volinsky (1999).
[21]See, e.g., Raftery and Zheng (2003) and Schervish (1995).

3.1 The explanatory power of the corporate bond factor zoo

We start by only considering the pricing power of the 25 traded and nontraded bond market factors as we are interested in gauging to what extent the cross-section of corporate bond returns is priced by bond market specific factors. The test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity and the 14 traded bond factors. We use the continuous spike-and-slab approach Section 2 and we report both posterior probabilities (given the data) of each factor (i.e., $\mathbb{E}[\gamma_j|\text{data}], \forall j$) and posterior means of the factors' price of risk (i.e., $\mathbb{E}[\lambda_j|\text{data}], \forall j$) in Figure 1 and Table 1. The prior for the Sharpe ratio achievable with the linear SDF are set to a range of values computed as 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors which is equal to 3 annualized. That is, from a very strong degree of shrinkage (20%, i.e., a prior Sharpe ratio value of 0.6 annualised), to a very moderate one (80% or a prior annualized Sharpe ratio of 2.4).

Recall that we have a uniform (hence flat) prior for the model dimensionality and each factor has an ex ante expected probability of being selected equal to 50% (dashed horizontal line in Figure 1). Figure 1 illustrates that—with some notable exceptions—most factors proposed in the corporate bond literature have a posterior probability of being part of the SDF that is below 50%.

Several observations are in order. First, given its posterior probabilities across the range of prior Sharpe ratios considered, there is strong evidence for including the PEADB (i.e., the bond post-earnings announcement drift) factor as a source of priced risk in the SDF. This is a rather surprising result, as PEADB has not specifically been proposed as a priced risk factor in the corporate bond literature. Nozawa, Qiu, and Xiong (2023) are the first to document a post-earnings announcement drift in corporate bond prices and they rationalise their findings in a stylised model of disagreement. They also show that a strategy that purchases bonds issued by firms with high earnings surprises and sells bonds of firms with low earnings surprises generates sizable Sharpe ratios and large risk-adjusted returns. On the other hand, Bryzgalova, Huang, and Julliard (2023) and Avramov, Cheng, Metzker, and Voigt (2023) find strong evidence that the *stock market* post-earnings announcement drift

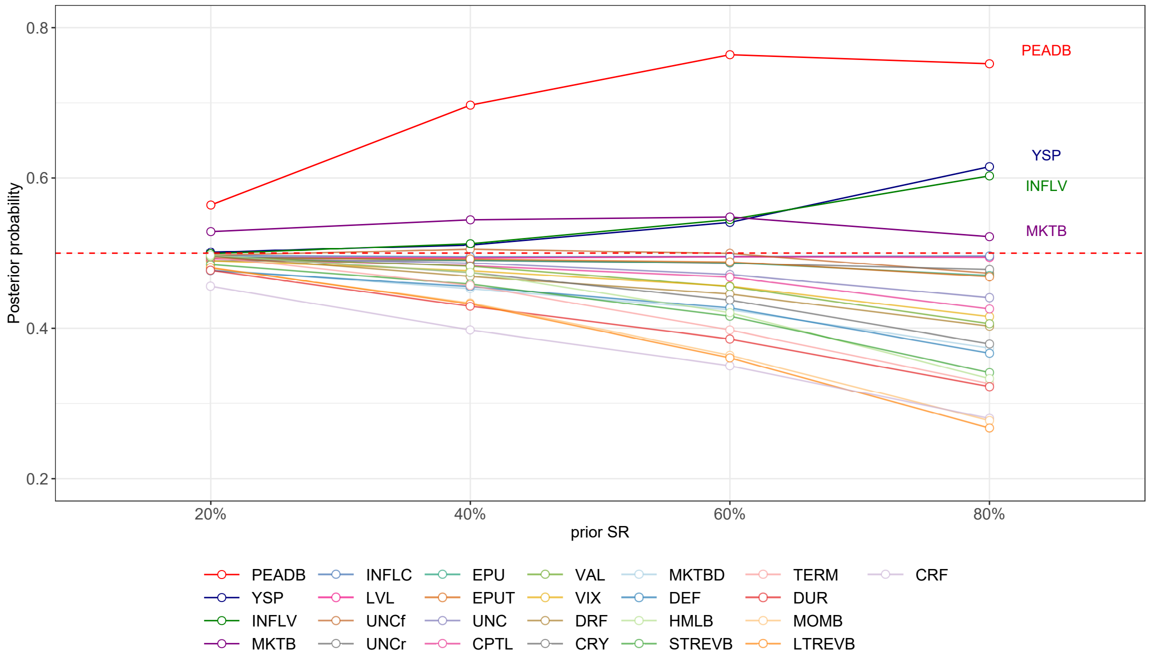


Figure 1: Posterior factor probabilities – bond factor zoo.

Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, of the 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. Posterior probabilities for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR_{\boldsymbol{f}}^2 \mid \sigma^2]}$, set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

(PEAD) factor of Daniel, Hirshleifer, and Sun (2020) exhibits a particularly strong posterior probability for being included in the SDF for equity returns. Thus, it may seem natural to expect a similar level of performance for the bond version of the factor. Second, the bond market factor (MKTB) exhibits posterior probabilities above 50% for the full range of prior Sharpe ratios, corroborating the findings in Dickerson, Mueller, and Robotti (2023) that the bond market factor carries valuable information to explain risk premia of corporate bonds. The result also mirrors the finding in Bryzgalova, Huang, and Julliard (2023) who conclude that the stock market factor is a source of priced risk for equity returns when considered in conjunction with all other factors in the equity factor zoo. Third, there is a small number of factors that have posterior probabilities above 50% percent for all but very low values

Table 1: Posterior factor probabilities and risk prices – bond specific factor zoo.

	Factor prob., $\mathbb{E}[\gamma_j \mid \text{data}]$				Price of risk, $\mathbb{E}[\lambda_j \mid \text{data}]$			
	Total prior Sharpe ratio				Total prior Sharpe ratio			
Factors:	20%	40%	60%	80%	20%	40%	60%	80%
PEADB	0.564	0.697	0.764	0.757	0.073	0.285	0.511	0.646
YSP	0.501	0.511	0.541	0.613	0.004	0.016	0.048	0.150
INFLV	0.499	0.513	0.545	0.604	0.006	0.026	0.072	0.199
MKTB	0.529	0.544	0.548	0.519	0.066	0.171	0.291	0.423
INFLC	0.497	0.495	0.496	0.495	0.000	0.000	0.000	0.001
LVL	0.495	0.494	0.495	0.496	0.000	−0.002	−0.004	−0.012
UNCf	0.497	0.505	0.500	0.475	−0.010	−0.035	−0.074	−0.133
UNCr	0.490	0.490	0.487	0.480	0.000	0.003	0.010	0.032
EPU	0.495	0.491	0.487	0.472	0.001	0.005	0.014	0.038
EPUT	0.493	0.492	0.488	0.467	0.000	0.001	−0.005	−0.018
UNC	0.494	0.487	0.471	0.439	−0.003	−0.009	−0.011	−0.001
CPTL	0.493	0.483	0.468	0.422	−0.006	−0.028	−0.058	−0.087
VAL	0.495	0.482	0.455	0.404	0.041	0.094	0.148	0.196
VIX	0.490	0.476	0.456	0.415	−0.008	−0.021	−0.049	−0.081
DRF	0.496	0.469	0.445	0.402	0.026	0.033	0.010	−0.049
CRY	0.497	0.474	0.437	0.378	0.040	0.083	0.115	0.140
MKTBD	0.478	0.453	0.424	0.377	0.019	0.016	−0.007	−0.054
DEF	0.476	0.455	0.427	0.369	−0.017	−0.071	−0.130	−0.17
HMLB	0.497	0.474	0.420	0.332	0.047	0.097	0.117	0.104
STREVB	0.485	0.459	0.416	0.343	0.001	−0.003	−0.009	−0.017
TERM	0.495	0.457	0.398	0.324	0.075	0.137	0.151	0.135
DUR	0.477	0.429	0.386	0.325	0.020	0.004	−0.035	−0.073
MOMB	0.480	0.434	0.364	0.276	−0.020	−0.032	−0.027	−0.014
LTREVB	0.480	0.432	0.361	0.267	0.021	0.036	0.030	0.018
CRF	0.456	0.398	0.350	0.280	0.009	0.036	0.074	0.095

Posterior probabilities, $\mathbb{E}[\gamma_j \mid \text{data}]$, and posterior mean of (annualized) risk prices, $\mathbb{E}[\lambda_j \mid \text{data}]$, of the 25 bond specific factors described in Appendix B. The prior for each factor inclusion is a Beta(1, 1), yielding a prior expectation for γ_j of 50%. Test assets are the returns, in excess of the one-month risk-free rate, of 50 bond portfolios sorted on credit spreads, size, rating and maturity, plus the 14 traded bond factors ($N = 64$). Results are tabulated for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, with values set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the test assets. Sample period: 1986:01 to 2021:09 ($T = 429$).

of prior Sharpe ratio. In particular, there is some evidence that the (nontradable) inflation volatility risk factor INFLV (see, e.g., Kang and Pflueger (2015) and Ceballos (2022)) and the slope of the yield curve YSP (see, e.g., Koijen, Lustig, and Van Nieuwerburgh 2017) should be included in the SDF.

That said, while there are a few factors for which the posterior probability is roughly equal to the prior (implying that at least some of these factors, as discussed in the next subsection, are likely to be weakly identified at best), there are a large set of factors that is unlikely to be part of the SDF pricing the cross-section of bond returns. Specifically, besides PEADB and MKTB, the traded bond market factors are overall *highly* unlikely to be individually included in the SDF. For instance, with a prior Sharpe ratio set to 80% of the ex post maximum, the posterior probabilities of LTREVB, MOMB, and STREVB range from 0.27 to 0.34.[22] Other traded bond factors related to default and credit risk (DRF and CRF) are also quite unlikely candidates to be included in the SDF. This finding is consistent with Dickerson, Mueller, and Robotti (2023) who show that among the low dimensional models they consider, other factors seem to be spanned by MKTB. However, as we show in Internet Appendix IA.2, albeit MKTB carries a sizable, and statistically significant, ex-post risk premium in two-pass regressions, the CAPMB—a factor model of the SDF with only MKTB—is not supported by the data.

As pointed out in van Binsbergen, Nozawa, and Schwert (2023), to correct for the effect of the secular decline in interest rates affecting most of our sample, it might be preferable to construct returns on the test assets in excess of a duration-matched government bond return, rather than using a simple risk free rate. The rationale for this duration-adjustment is to isolate the portion of a bond performance that is ascribable solely to the credit risk of each bond. Furthermore, they show that not only does the duration adjustment significantly improve the ability of the CAPM to price corporate bonds, but it also changes the performance of the factor models considered in the previous literature. Therefore, as explained in Appendix C, we perform the duration adjustment for all our test assets and tradable factors, and re-estimate both factor posterior probabilities and market prices of risk. As shown in Figure A.1 and Table A.2 of Appendix C, we obtain almost identical results using these duration-adjusted returns.

Overall, we conclude that there are only very few bond specific factors that, given the data, are likely components of the SDF, and that only one factor—the bond post-earnings announcement drift factor (PEADB)—is clearly a robust source of priced risk in the cross-section of corporate bond returns. Next, we expand the set of factors to include those proposed in the equity literature.

[22]See Table A.1 in Appendix B for a detailed description of the factors.

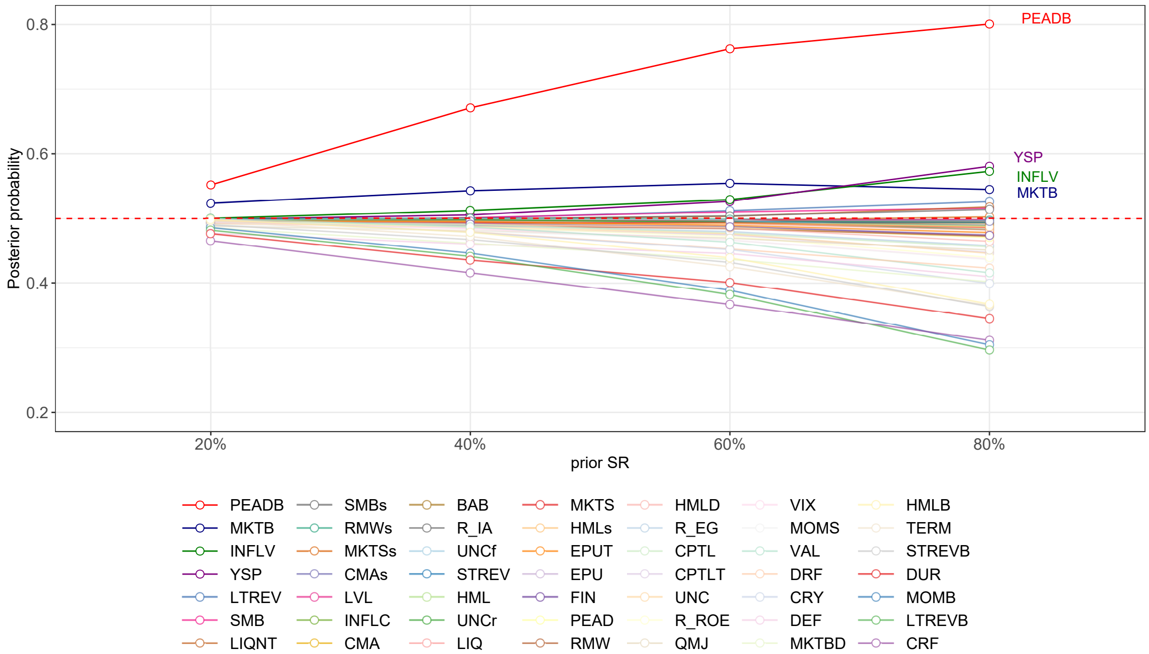


Figure 2: Posterior factor probabilities – bond and stock factor zoo.

Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, of 49 bond and stock factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. Posterior probabilities for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR_{\boldsymbol{f}}^2 \mid \sigma^2]}$, set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

3.2 Pricing corporate bonds with equity market information

We now include an additional 24 factors from the equity factor zoo, bringing the total number of factors that we consider to 49 and, consequently, yielding a space of about 563 trillion models to consider. As we are only interested in eliciting which factors contain pricing information for the cross-section of corporate bond returns, we do not add the traded equity factors to the cross-section of test assets. We proceed as before and report both posterior probabilities (given the data) of each factor (i.e., $\mathbb{E}[\gamma_j|\text{data}], \forall j$) and posterior means of the factors' price of risk (i.e., $\mathbb{E}[\lambda_j|\text{data}], \forall j$) in Figure 2 and Table 2.

In terms of posterior probabilities, the set of most likely four factors does not change

Table 2: Posterior factor probabilities and risk prices – bond and stock factor zoo

	Factor prob., $\mathbb{E}[\gamma_j\|\text{data}]$				Price of risk, $\mathbb{E}[\lambda_j\|\text{data}]$			
	Total prior Sharpe ratio				Total prior Sharpe ratio			
Factors:	20%	40%	60%	80%	20%	40%	60%	80%
PEADB	0.552	0.671	0.762	0.801	0.055	0.228	0.463	0.685
MKTB	0.523	0.543	0.554	0.545	0.053	0.144	0.260	0.413
INFLV	0.500	0.512	0.529	0.573	0.005	0.020	0.052	0.137
YSP	0.498	0.505	0.526	0.581	0.003	0.012	0.034	0.101
LTREV	0.493	0.500	0.512	0.526	0.005	0.023	0.064	0.151
SMB	0.494	0.501	0.510	0.516	0.008	0.033	0.088	0.185
LIQNT	0.496	0.498	0.503	0.518	−0.002	−0.008	−0.022	−0.062
SMBs	0.495	0.499	0.504	0.514	0.002	0.010	0.026	0.064
RMWs	0.499	0.500	0.499	0.497	0.000	0.000	0.000	0.000
MKTSs	0.497	0.498	0.497	0.502	−0.002	−0.009	−0.025	−0.06
CMAs	0.498	0.496	0.496	0.495	0.000	0.001	0.002	0.004
LVL	0.495	0.493	0.497	0.498	0.000	−0.001	−0.003	−0.008
INFLC	0.496	0.495	0.496	0.494	0.000	0.000	0.000	0.001
CMA	0.496	0.497	0.493	0.494	0.000	0.001	0.001	−0.002
BAB	0.494	0.497	0.494	0.492	0.002	0.008	0.020	0.045
R_IA	0.497	0.497	0.494	0.489	0.000	0.000	−0.001	−0.005
UNCf	0.496	0.499	0.497	0.482	−0.008	−0.028	−0.059	−0.104
STREV	0.496	0.491	0.494	0.494	0.003	0.012	0.035	0.092
HML	0.496	0.496	0.494	0.488	0.000	0.001	0.002	0.002
UNCr	0.496	0.498	0.493	0.483	0.000	0.002	0.007	0.021
LIQ	0.496	0.495	0.492	0.482	−0.001	−0.005	−0.012	−0.023
MKTS	0.493	0.492	0.493	0.486	−0.006	−0.032	−0.081	−0.165
HMLs	0.492	0.491	0.492	0.487	0.000	0.001	0.003	0.006
EPUT	0.496	0.495	0.489	0.478	0.000	0.000	−0.003	−0.019
EPU	0.495	0.492	0.487	0.475	0.001	0.004	0.011	0.031
FIN	0.493	0.491	0.487	0.474	0.000	0.004	0.011	0.022
PEAD	0.494	0.491	0.486	0.470	−0.001	−0.002	−0.005	−0.008
RMW	0.494	0.489	0.484	0.473	0.001	0.006	0.016	0.038
HMLD	0.494	0.490	0.485	0.465	0.006	0.021	0.049	0.088
R_EG	0.492	0.489	0.480	0.458	0.001	0.005	0.008	0.018
CPTL	0.494	0.491	0.477	0.456	−0.005	−0.022	−0.044	−0.07
CPTLT	0.495	0.485	0.475	0.447	−0.003	−0.018	−0.035	−0.043
UNC	0.492	0.487	0.476	0.446	−0.002	−0.007	−0.01	−0.005
R_ROE	0.496	0.487	0.472	0.440	−0.006	−0.015	−0.028	−0.044
QMJ	0.488	0.481	0.469	0.451	−0.008	−0.021	−0.05	−0.111
VIX	0.493	0.484	0.466	0.437	−0.006	−0.017	−0.039	−0.073
MOMS	0.490	0.481	0.466	0.436	−0.003	−0.005	−0.004	0.007
VAL	0.500	0.487	0.463	0.416	0.033	0.079	0.124	0.167
DRF	0.497	0.479	0.453	0.423	0.022	0.034	0.02	−0.024
CRY	0.499	0.478	0.452	0.400	0.033	0.071	0.103	0.132
DEF	0.477	0.460	0.446	0.410	−0.012	−0.058	−0.122	−0.19
MKTBD	0.485	0.462	0.437	0.402	0.016	0.017	−0.005	−0.061
HMLB	0.498	0.479	0.439	0.368	0.038	0.086	0.115	0.117
TERM	0.496	0.472	0.425	0.366	0.062	0.13	0.161	0.168
STREVB	0.490	0.467	0.432	0.364	0.001	−0.002	−0.01	−0.021
DUR	0.476	0.436	0.401	0.345	0.018	0.013	−0.015	−0.044
MOMB	0.486	0.446	0.389	0.305	−0.016	−0.03	−0.027	−0.017
LTREVB	0.482	0.441	0.382	0.297	0.017	0.031	0.023	−0.001
CRF	0.465	0.416	0.367	0.312	0.006	0.027	0.061	0.093

Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, and posterior mean of (annualized) risk prices, $\mathbb{E}[\lambda_j|\text{data}]$, of the 49 bond and equity factors described in Appendix B. The prior for each factor inclusion is a Beta(1, 1), yielding a prior expectation for γ_j of 50%. Test assets are the returns, in excess of the one-month risk-free rate, of 50 bond portfolios sorted on credit spreads, size, rating and maturity, plus the 14 traded bond factors ($N = 64$). Results are tabulated for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, with values set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the test assets. Sample period: 1986:01 to 2021:09 ($T = 429$).

when the equity factors are included. PEADB still has the highest posterior probability, significatly larger than its prior value, and, thus, there is strong evidence to include the post-earnings announcement drift for bonds in the SDF. The post-earnings announcement drift for stocks instead (PEAD), that should be part of the SDF to price the cross-section of equity returns according to previous studies, exhibits posterior probabilities below 50% for the full range of prior Sharpe ratios. Apart from PEADB, the factors with the highest posterior probabilities are still MKTB, INFLV and YSP.

The equity factors that have posterior probabilities above 50% are the long-term reversal factor LTREV of Jegadeesh and Titman (2001),[23] the SMB factor of Fama and French (1992), the liquidity factor LIQNT of Pástor and Stambaugh (2003),[24] and the SMBs factor of Daniel, Mota, Rottke, and Santos (2020) (which is the SMB factor of Fama and French (1993) without its unpriced component). While most posterior probabilities for these factors are *individually* below 53%, the results in Sections 3.3 and 3.4 below stress the importance of including stock market information to construct the BMA-SDF, i.e., the results suggest that equity-based factors are *jointly* useful to price the cross-section of corporate bond returns since, as we are about to show, the "true" latent SDF appears to be *dense* in the space of observable factors.

Finally, as illustrated in Figure A.2 and Table A.3 of Appendix C, results are very stable when considering duration-adjusted corporate bond returns.[25]

3.3 Which, and how many, factors are needed to price corporate bonds?

The results in the previous sections highlight that in the joint zoo of equity and bond factors only one variable—PEADB— should be included in an empirical SDF to price corporate bond returns with a very high probability. As we show in Section 3.4 below, PEADB

[23]The long-term reversal factor is constructed as a long-short portfolio of stocks sorted on their cumulative return accrued in the previous 60-13 months.

[24]This nontradable liquidity factor is computed as the average of individual-stock measures estimated as the daily data residual predictability after controlling for the market factor

[25]The posterior probabilities are also very similar over the shorter sample (2002:09-2021:09) when using the WRDS TRACE data reported in Figure IA.2 in the Internet Appendix.

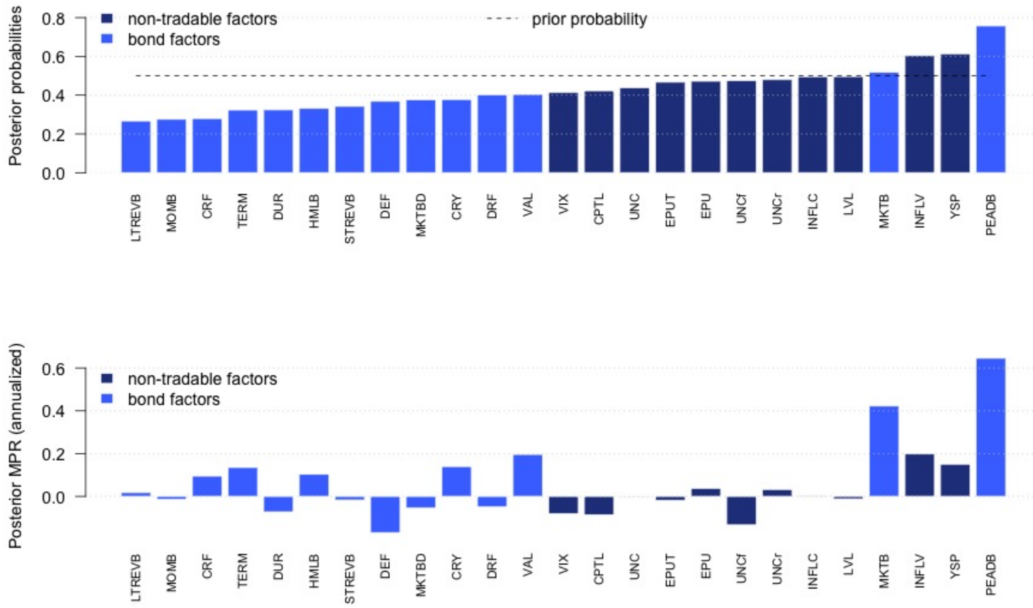


Figure 3: Posterior factor probabilities and market prices of risk – bond factor zoo.

Posterior factor probabilities (top panel), $\mathbb{E}[\gamma_j|\text{data}]$, and the corresponding posterior market prices of risk (bottom panel), $\mathbb{E}[\lambda_j|\text{data}]$, of 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

alone is not sufficient to ensure good pricing performance, and several other factors should be included in the model averaging. Before we turn to the cross-sectional pricing performance of various models against the BMA-SDF, we first analyze the nature of the various factors, and the dimensionality of the most likely SDF given the data.

Since the results over a wide range of prior values for the Sharpe ratio in Sections 3.1 and 3.2 appear very stable for factor selection, we now focus on just one value for this quantity, as this tends to deliver the best out-of-sample performance (discussed in the next section): a prior equal to 80% of the ex post maximum Sharpe ratio. Very similar results are obtained for different values of the prior Sharpe ratio and are reported in Appendix D.

Figure 3 reports the posterior probabilities (top panel) and the posterior (annualized) market price of risk (bottom panel) of the bond specific factors in the estimation *without* equity market factors. Note that out of the four factors with posterior probabilities higher than their prior values (i.e., PEADB, YSP, INFLV and MKTB), two of them are nontradable in nature and, importantly, all four command substantial market prices of risk. Furthermore, the next nine factors with highest posterior probabilities are also all nontradable in nature, and many of them command sizable market prices of risk. Moreover, the posterior median number of nontradable factors in the "true" latent SDF is 5 (with a centered 95% coverage between 2 and 9), and the posterior probability of zero nontradable factors in the SDF is about 0.05%. Nevertheless, the risk prices of two of these nontradable factors (namely INFLC and UNC), are effectively shrunk to zero. This is due to the fact that these are likely *weak factors* in the cross-section of corporate bonds.[26] The occurrence of weak factors, which, in fact, is most common among the nontradable ones, causes identification failure and invalidates canonical estimation approaches (e.g., GMM, MLE and two-pass regressions). This is not an issue for our method, which restores inference, by design, by effectively shrinking their market prices of risk towards zero. Finally, and possibly surprisingly, the remaining 12 tradable factors, that have been specifically designed to price the cross-section of corporate bond returns, are all less likely to be included in the SDF, with posterior probabilities substantially lower than their prior values.

Overall, these findings suggest that i) the bond market specific tradable factors suggested in the previous literature (with the exception of MKTB) are not likely sources of priced risk and that ii) nontradable factors are—if not necessarily individually (with the exception of YSP and INFLV)—at least jointly useful sources of information for pricing corporate bond returns. This feature is further stressed by Figure 4 that reports the posterior dimension of the SDF (top panel) in terms of observable bond specific factors to be included in it, and the posterior distribution of the Sharpe ratios achievable with such an SDF (bottom panel). As shown in the top panel, a substantial number of corporate bond factors (10 or 11 on average, with a centered 95% coverage of 6 to 16 factors) are needed to construct a likely SDF out of the corporate bond factor zoo, implying that the low dimensional models suggested in

[26]That is, their correlations with the test assets are small and have little cross-sectional dispersion, especially in the case of INFLC. See, e.g., Gospodinov, Kan, and Robotti (2019), Kleibergen (2009), and Bryzgalova, Huang, and Julliard (2023) for a formal definition for weak and level factors.

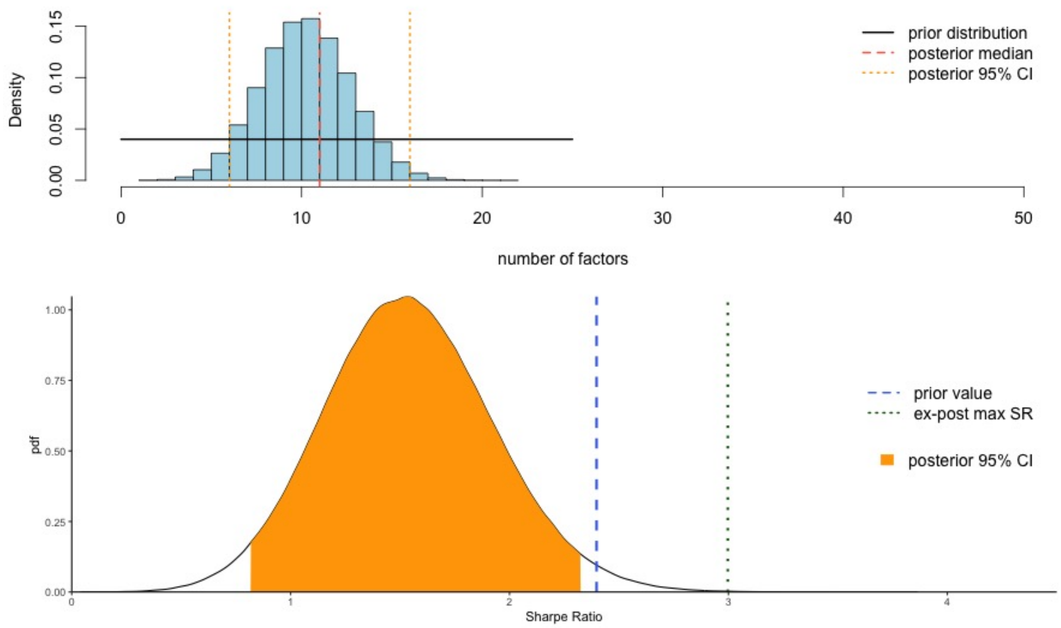


Figure 4: Posterior SDF dimension and Sharpe ratio – bond factor zoo.

Posterior distributions of the number of factors to be included in the SDF (top panel) and of the SDF-implied Sharpe ratio (bottom panel), computed using 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

the previous literature have very weak support in the data and are misspecified with high probability. As shown in Figures IA.3 and IA.4 of the Internet Appendix, these findings are robust to the particular value of the prior Sharpe ratio (albeit, as expected, less informative when we impose an extreme degree of shrinkage).

Given that the previous corporate bond literature has exclusively focused on low dimensional factor models, one might worry that our uniform prior for the model dimensionality might be the reason behind the evidence in support of nontradable, and against tradable (with the exception of PEADB and, to a lesser extent, MKTB), factors as likely drivers of the SDF. Fortunately, our Bayesian approach allows to encode a strong prior belief in support of sparse models. In particular, we recompute factor posterior probabilities using a Beta$(3, 12)$ as prior for the individual factors inclusion. This tightly imposes a prior in favour of sparsity: the a priori expected number of factors is 5, with a prior median of 3, and vanishing probabilities for dense models, yielding a posterior distribution for the dimension of the models that ranges from 1 to 7–9 factors (with the upper bound depending on the prior Sharpe ratio value).

As shown in Figure A.8 of Appendix D, even when imposing sparsity, our results are quite stable: nontradable factors are jointly important for pricing corporate bonds and only PEADB, INFLV and YSP have posterior probabilities above the prior value (equal to 0.20 in this case), while the posterior probability of MKTB also exceeds its prior value when the prior Sharpe ratio is set to 60% or less. This last result might seem somehow surprising given the findings in Dickerson, Mueller, and Robotti (2023) that the CAPMB is not dominated by the sparse factor models considered in their study. But this difference is due to three simple reasons. First, our estimation method considers many additional sparse models that were *not* in the previous literature, and many of these models (especially when including PEADB, INFLV and YSP) have better pricing ability than the simple CAPMB model. Second, Dickerson, Mueller, and Robotti (2023) use a method designed for *tradable* factors, and apply it to the nontradable ones by first projecting them on the space of tradable factors (not including PEADB). Hence, the ability of tradable factors to matter for pricing is limited by the pricing ability of the tradable factors they are projected on—and, as shown above and also in the next section, most corporate bond tradable factors have very limited pricing ability. Third, as shown in Internet Appendix IA.2, an SDF with only MKTB as the sole driving source of priced risk is *not* supported by the data. Nevertheless, our findings show that MKTB carries relevant information for pricing corporate bonds, but simply not enough for doing so in isolation.

Next, we turn to examining whether equity-based factors contain valuable information to price corporate bonds. As shown in the top panel of Figure 5, equity-based factors are overall more likely components of the true latent SDF than bond tradable factors (except PEADB and MKTB). Hence, we would expect them to play a non-trivial role in the BMA-SDF. Moreover, as shown in the bottom panel of Figure 5, they often demand sizable market prices of risk in the corporate bond space (while some of them appear to be weak factors in the

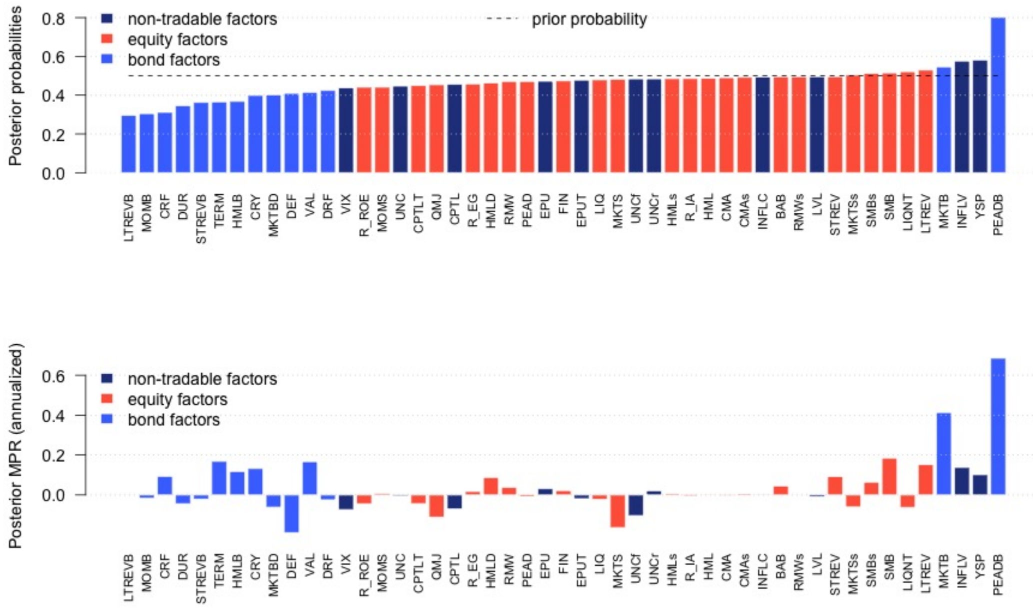


Figure 5: Posterior factor probabilities and market prices of risk – bond and equity factor zoo.

Posterior factor probabilities (top panel), $\mathbb{E}[\gamma_j|\text{data}]$, and the corresponding posterior market prices of risk (bottom panel), $\mathbb{E}[\lambda_j|\text{data}]$, of 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

space of corporate bonds, and their risk prices are effectively shrunk to zero). Furthermore, this outperformance of equity factors relative to the bond specific tradable ones holds true for different values of the Sharpe ratio prior (Figure IA.5 of the Internet Appendix), and even when we impose sparcity of the SDF (Figure A.9 of Appendix D).

The top panel of Figure 6 also reveals that the posterior dimension of the SDF becomes much larger once we allow for the potential inclusion of equity-based factors: the posterior median of the number of factors in the SDF is now 23, and the centered 95% coverage area ranges from 16 to 29 factors. This implies that the low-dimensional factor models considered in the previous literature are misspecified with extremely high probability. Note that this finding, as shown in Figure IA.3 of the Internet Appendix, is stable for all values of prior

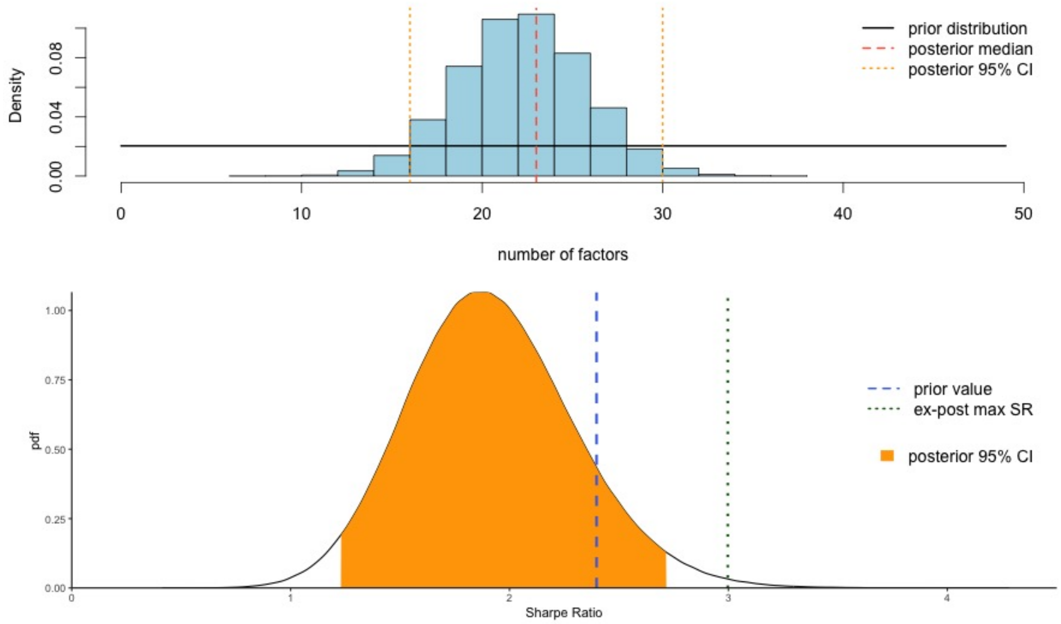


Figure 6: Posterior SDF dimension and Sharpe Ratio – bond and equity factor zoo.

Posterior distributions of the number of factors to be included in the SDF (top panel) and of the SDF-implied Sharpe ratio (bottom panel), computed using the 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

Sharpe ratio. Furthermore, the posterior median number of equity-based factors in the SDF is 12 (with a centered 95% coverage between 7 and 16), and the posterior probability of no such factor in the SDF is of the order of 10^{-9}. That is, equity market information is salient for pricing corporate bond returns.

Interestingly, as shown in the bottom panel of Figure 6, once we allow the SDF to potentially load on equity factors, the ex post maximum Sharpe ratio achievable in the data with the 64 bond portfolios and traded factors is no more unrealistically large relative to the posterior distribution of the SDF-implied Sharpe ratio. This is in sharp contrast with what we observe in the bottom panel of Figure 4, where the ex post maximum Sharpe ratio of the data has zero probability to be generated by an SDF that comprises only bond specific

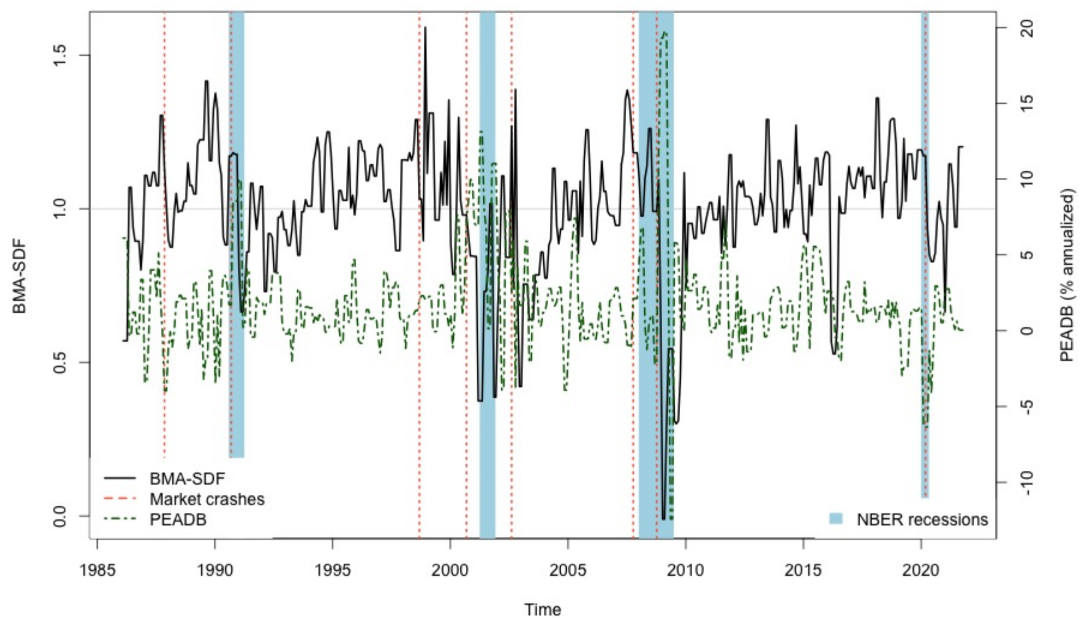


Figure 7: BMA-SDF, economic cycles, and PEADB – bond and equity factor zoo.

Smoothed time series of the posterior mean of the BMA-SDF (left scale), computed using 49 bond and equity factors described in Appendix B, and of the smoothed PEAD factor annualized returns (right scale). The blue shaded areas represent NBER-dated recessions, and the red dotted vertical lines correspond to the major stock market crashes identified in Mishkin and White (2002) plus the 2008 and 2020 contractions. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

factors. Once again, this indicates that the bond factor zoo needs complementing with factors from the equity literature to produce a likely SDF for the corporate bond market.

Notably, as shown in Figure A.3 of Appendix C, the importance of nontradable factors, and the poor performance of tradable ones (except PEADB and, to a much lower extent, MKTB) is confirmed when focusing on duration-adjusted corporate bond returns. Furthermore, even when considering these test assets that are meant to isolate and focus solely on the credit risk component of bond returns, the SDF appears to be dense in the space of observable factors (as shown in Figure A.4 of Appendix C). That is, even duration-adjusted corporate bond returns are extremely unlikely to have been generated by a low dimensional

SDF as the ones considered in the previous literature.

The important contribution of nontradable and equity-based factors to the BMA-SDF can also be elicited from Figure 7 that plots the (posterior mean of the) BMA-SDF and the return series of the PEADB factor. The BMA-SDF has a clear business cycle pattern: it increases during expansions and peaks right before recessions and around the time of financial market crashes. Hence, our approach yields an estimated SDF that behaves as one would expect from the intertemporal marginal rate of substitution of an agent that prices asset returns and is exposed to the risk arising from the general economic conditions. Furthermore, its close links to business cycles and market crashes is in line with the SDFs estimated in the literature using very different test assets and econometric methods (see, e.g., Ghosh, Julliard, and Taylor 2016). The PEADB return series instead, that is (by construction, given its high posterior probability of being part of the SDF) highly correlated with the BMA-SDF (the posterior mean correlation is about -0.51), does not have such a clear business cycle pattern and has a weaker association with market-wide crashes. That is, the additional factors included in the BMA-SDF are important in determining these economically salient features of its time series behaviour.

3.4 Cross-sectional asset pricing

We now focus on the cross-sectional asset pricing performance of our BMA estimates of the Stochastic Discount Factor (BMA-SDF), constructed with and without the inclusion of equity-based factors, both in- and out-of-sample, and compare it with both traditional popular reduced-form factor models for corporate bonds, and low dimensional models constructed by selecting only the most likely factors (as suggested, e.g., in Barillas and Shanken 2018). Specifically, we consider two “top factor” models using (i) bond factors and (ii) a combination of bond and equity factors based on the analysis in Sections 3.1 and 3.2. The bond model includes PEAD, MKTB, INFLV and YSP, or the 4 factors with a posterior probability above 50% for a wide range or prior Sharpe ratios. Allowing for all factors, we add to the above LTREV, LIQNT, SMB and SMBs, who all exhibit posterior probabilities for inclusion in the SDF above 50%, yielding an 8-factor model. Table 3 reports root mean squared pricing error (RMSE), mean absolute pricing errors (MAPE), and OLS and GLS cross-sectional R^2 for a variety of models and test assets.[27] The test assets for the in-sample tests are the 50 bond portfolios and the 14 traded bond factors, whereas the cross-sectional out-of sample tests are conducted using a different set of 77 bond portfolios.[28] In the cross-sectional out-of-sample tests, the SDFs are first estimated using the baseline in-sample test assets and then used to price (without additional parameter estimation) the out-of-sample test assets.

Panel A reports the statistics for the BMA-SDF constructed using only the 25 bond factors (for a total of 33.6 million models) whereas Panel B contains the results for the BMA-SDF based on 49 bond and equity factors (for a total of about 563 trillion models included in the model averaging). For a benchmark comparison, in Panel C we consider the bond CAPM (CAPMB), the equity CAPM, and the original Fama and French (1993) five-factor model (FF5), which includes the market return (MKTS), a proxy for the size effect (SMB), a proxy for the book-to-market anomaly (HML), a proxy for the default risk of bonds (DEF), and the slope of the term structure of Government bonds (TERM). In addition, we include a single-factor model with PEADB (which has the highest posterior probability of being included in the SDF) and the two "top factor" models described above. Results for the (GLS) BMA-SDFs are reported for a range of prior values of the Sharpe ratio achievable in the economy. All the benchmark model SDFs are estimated via a GLS version of the GMM (see, e.g., Cochrane (2005, pp. 256–258)). Note that for the cross-sectional out-of-sample pricing we do not refit the BMA-SDF, nor the other benchmark models, to the new data. Instead, we use the estimated parameters from the respective in-sample pricing exercises.

The results for the BMA-SDF based on bond factors only are presented in Panel A of Table 3. Over the range of Sharpe ratio priors, it is clear that a relatively high prior Sharpe ratio produces the best results. Both in- and out-of-sample, the BMA-SDF strongly outperforms all the low dimensional models proposed in the previous literature (Panel C),

[27] All data is normalised by the standard deviations, hence pricing errors are expressed in (annualised) Sharpe ratio units. The measures of fit are computed as: $RMSE \equiv \sqrt{\frac{1}{N}\sum_{i=1}^{N}\alpha_i^2}$, $MAPE \equiv \frac{1}{N}\sum_{i=1}^{N}|\alpha_i|$, $R^2_{ols} \equiv 1 - \frac{\left(\alpha-\frac{1}{N}\alpha^\top 1_N\right)^\top\left(\alpha-\frac{1}{N}\alpha^\top 1_N\right)}{\left(\alpha-\frac{1}{N}\alpha^\top 1_N\right)^\top\left(\mu_R-\frac{1}{N}\mu^\top R 1_N\right)^\top\left(\mu_R-\frac{1}{N}\mu^\top R 1_N\right)}$ and $R^2_{gls} \equiv 1 - \alpha^\top \Sigma_R^{-1}\alpha\mu^\top R\left(\Sigma_R^{-1}\right)R\mu_R$

[28] The out-of-sample portfolios include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the 17 Fama French industry portfolios discussed in Section 1.

Table 3: Cross-sectional asset pricing.

	In-sample				Out-of-sample			
	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}
Panel A: BMA-SDF with 25 bond factors (33.6 mn models)								
prior SR = 20%	0.192	0.144	0.240	0.125	0.127	0.091	0.088	0.035
prior SR = 40%	0.163	0.127	0.450	0.194	0.119	0.087	0.197	0.060
prior SR = 60%	0.144	0.112	0.574	0.252	0.113	0.085	0.272	0.075
prior SR = 80%	0.130	0.100	0.651	0.308	0.108	0.083	0.336	0.092
Panel B: BMA-SDF with 49 bond and stock factors (563 tn models)								
prior SR = 20%	0.198	0.147	0.195	0.117	0.128	0.092	0.068	0.031
prior SR = 40%	0.171	0.132	0.399	0.187	0.121	0.087	0.176	0.062
prior SR = 60%	0.150	0.117	0.539	0.261	0.114	0.083	0.272	0.086
prior SR = 80%	0.132	0.102	0.641	0.348	0.106	0.080	0.368	0.112
Panel C: Benchmark models and most likely factors								
CAPMB	0.176	0.115	0.365	0.137	0.125	0.094	0.113	0.056
CAPM	0.261	0.207	−0.401	0.032	0.156	0.110	−0.368	−0.023
FF5	0.205	0.147	0.138	0.180	0.142	0.106	−0.146	0.011
PEADB	0.255	0.173	−0.337	0.156	0.118	0.087	0.217	−0.027
Top factors bond	0.135	0.095	0.627	0.490	0.122	0.104	0.154	0.026
Top factors all	0.136	0.099	0.618	0.574	0.111	0.091	0.307	0.097

In-sample and cross-sectional out-of-sample pricing performance of BMA-SDF, notable factor models, and factors with a posterior probability greater than 50%. We use GMM-GLS to estimate factor risk prices for CAPMB, CAPM, and the and Fama and French (1992, 1993) model, which includes the MKTS, SMB, HML, DEF and TERM factors, and a single-factor model with PEADB. The 'Top factors bond' model includes PEADB, MKTB, INFLV and YSP. The 'Top factors all' model includes additionally LTREV, LIQNT, SMB and SMBs. For the BMA-SDF, we report results for a range of prior Sharpe ratio values that are set as 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. In-sample (IS) test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity and the 14 traded bond factors ($N = 64$). Out-of-sample (OS) test assets assets include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the 17 Fama French industry portfolios ($N = 77$). In cross-sectional OS tests, models are first estimated using the baseline IS test assets and then used to price (with no additional parameter estimation) the OS assets. All data is standardized, that is, pricing errors are in Sharpe ratio units.

except when we impose an extreme degree of prior shrinkage (a prior Sharpe ratio equal to 20% of the ex post maximum Sharpe ratio of the test assets, i.e. a prior for the Sharpe achievable with all the factors of only 0.6 per annum), in which case CAPMB has marginally better in-sample, and similar out-of-sample, pricing ability.

Interestingly, the most likely factor models (bottom two rows of Panel C) also strongly outperform the benchmark models both in- and out-of-sample, but do not perform as well as the BMA-SDF out-of-sample. Furthermore, the top factor model that includes *both* equity and bond factors performs significantly better out-of-sample than the specification with only bond factors, and only this specification prices assets out of sample similarly to the bond-factors-only BMA-SDF with a prior Sharpe ratio of 80%. Note that PEADB features prominently in these two models, as well as, given its high posterior probability, in the BMA-SDF. Nevertheless, PEADB alone (forth row of Panel C) is not enough to construct an SDF that prices assets accurately in- or out-of-sample, consistent with our evidence in the previous section that dense models are strongly preferred by the data.

In Panel B, we use all 49 of our bond (25) and equity (24) factors to construct the BMA-SDF. The in-sample asset pricing performance of the BMA-SDF remains roughly the same for the MSE, MAPE and R^2_{OLS} and marginally improves for the R^2_{GLS}. The two top factor models again outperform the BMA-SDF in terms of R^2_{GLS} in-sample (yet not in terms of R^2 and RMSE). But interestingly, out-of-sample the BMA-SDF (with moderate to low shrinkage) outperforms these two specifications.[29]

To further illustrate the stability of the cross-sectional out-of-sample pricing of the BMA-SDF, we use the 7 component portfolios sets (decile portfolios sorted on bond historical value-at-risk, duration, bond value, bond book-to-market, momentum, long-term reversals and the 17 Fama French industry portfolios), that comprise our 77 out-of-sample test assets, and construct $2^7 - 1 = 127$ possible combinations of these asset sorts (containing between 1 and 7 sets). We then repeat the cross-sectional out-of-sample asset pricing exercise using each of these 127 combinations and report the average values and standard deviations (across samples) of RMSE, MAPE, R^2_{OLS} and R^2_{GLS} in Table 4.

Two insights are readily apparent from the average asset pricing metrics from pricing the 127 portfolio combinations. First, for all but heavy levels of shrinkage, the BMA-SDF is superior compared to the benchmark models, with *much* lower standard deviations (in square brackets). This holds both for the BMA-SDF based on bond factors only as well as for the BMA-SDF that allows for all factors. That is, not only does the BMA-SDF have better OSS pricing ability for the "average" cross-section, but also has a more stable performance across all the possible OSS cross-sections. For instance, the CAPMB (CAPM), generates an average R^2_{OLS} of 0.09 (-0.35) with 0.24 (0.24) standard deviation, whereas

[29]We report similar asset pricing results for the shortened WRDS TRACE sample over the 2002:09–2021:09 sample period in Table IA.I of the Internet Appendix.

Table 4: Out-of-sample cross-sectional asset pricing with 127 cross-sections.

	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}
Panel A: BMA-SDF with 25 bond factors (33.6 mn models)				
prior SR = 20%	0.122	0.091	0.075	0.109
	[0.021]	[0.015]	[0.085]	[0.075]
prior SR = 40%	0.116	0.086	0.170	0.159
	[0.020]	[0.015]	[0.125]	[0.098]
prior SR = 60%	0.111	0.084	0.238	0.180
	[0.018]	[0.014]	[0.123]	[0.105]
prior SR = 80%	0.106	0.082	0.299	0.202
	[0.015]	[0.012]	[0.120]	[0.110]
Panel B: BMA-SDF with 49 bond and stock factors (563 tn models)				
prior SR = 20%	0.124	0.092	0.059	0.098
	[0.020]	[0.015]	[0.072]	[0.068]
prior SR = 40%	0.117	0.087	0.151	0.158
	[0.020]	[0.015]	[0.119]	[0.093]
prior SR = 60%	0.110	0.083	0.238	0.193
	[0.018]	[0.014]	[0.124]	[0.101]
prior SR = 80%	0.103	0.080	0.330	0.229
	[0.015]	[0.011]	[0.120]	[0.107]
Panel C: Benchmark models and most likely factors				
CAPMB	0.121	0.091	0.090	0.188
	[0.025]	[0.022]	[0.242]	[0.131]
CAPM	0.148	0.110	−0.345	0.035
	[0.031]	[0.024]	[0.244]	[0.109]
FF5	0.137	0.104	−0.225	0.062
	[0.025]	[0.020]	[0.476]	[0.108]
PEADB	0.114	0.086	0.159	−0.085
	[0.014]	[0.011]	[0.242]	[0.070]
Top factors bond	0.120	0.102	0.040	0.121
	[0.013]	[0.013]	[0.389]	[0.160]
Top factors all	0.109	0.090	0.193	0.234
	[0.013]	[0.012]	[0.338]	[0.169]

Average out-of-sample asset pricing performance metrics, and associated standard deviations across cross-sections, of the BMA-SDF, notable factor models, and factors with a posterior probability greater than 50%. The metrics are averaged over 127 possible combinations of the 7 sets of out-of-sample test assets. The out-of-sample (OS) portfolios which form the combinations, include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the Fama French 17 industry portfolios. We use GMM-GLS to estimate factor prices of risk for the CAPMB (bond CAPMB), the equity CAPM, the original FF5 model of Fama and French (1992) and Fama and French (1993), which includes the MKTS, SMB, HML and the DEF and TERM factors, and a single-factor model with PEADB. The 'Top factors bond' includes PEADB, MKTB, INFLV and YSP. The 'Top factors all' includes the above as well as LTREV, LIQNT, SMB and SMBs. For the BMA-SDF, we report results for a range prior Sharpe ratio values that are set as 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. In the cross-sectional OS tests, the models are first estimated using the baseline IS test assets and then used to price (without additional parameters estimation), the OS assets. All data is standardized, that is, pricing errors are in Sharpe ratio units.

Table 5: Cross-sectional asset pricing – duration adjusted returns

	In-sample				Out-of-sample			
	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}
Panel A: BMA-SDF with 25 bond factors (33.6 mn models)								
prior SR = 20%	0.137	0.085	0.142	0.118	0.091	0.066	0.129	0.006
prior SR = 40%	0.118	0.083	0.362	0.205	0.089	0.066	0.165	0.011
prior SR = 60%	0.108	0.083	0.467	0.271	0.092	0.068	0.114	0.005
prior SR = 80%	0.103	0.082	0.513	0.322	0.093	0.069	0.089	0.009
Panel B: BMA-SDF with 49 bond and stock factors (563 tn models)								
prior SR = 20%	0.140	0.085	0.110	0.108	0.092	0.066	0.114	0.006
prior SR = 40%	0.122	0.081	0.322	0.189	0.088	0.065	0.177	0.017
prior SR = 60%	0.109	0.081	0.463	0.266	0.089	0.067	0.156	0.022
prior SR = 80%	0.100	0.078	0.542	0.338	0.090	0.067	0.141	0.038
Panel C: Benchmark models and most likely factors								
CAPMB	0.170	0.104	−0.318	0.074	0.098	0.072	−0.006	−0.002
CAPM	0.160	0.098	−0.162	0.066	0.099	0.074	−0.042	−0.015
FF5	0.156	0.102	−0.114	0.132	0.104	0.075	−0.145	−0.037
PEADB	0.186	0.134	−0.580	0.153	0.098	0.072	−0.005	−0.086
Top factors bond	0.157	0.122	−0.125	0.458	0.116	0.097	−0.433	−0.152
Top factors all	0.135	0.110	0.165	0.511	0.131	0.106	−0.804	−0.083

In-sample and cross-sectional out-of-sample pricing performance of BMA-SDF, notable factor models, and factors with a posterior probability greater than 50%. Test assets and traded bond factors are computed with returns in excess of the duration-matched U.S. Treasury Bond rate of return. We use GMM-GLS to estimate factor risk prices for CAPMB, CAPM, and the and Fama and French (1992, 1993) model, which includes the MKTS, SMB, HML, DEF and TERM factors, and a single-factor model with PEADB. 'Top factors bond' includes PEADB, MKTB, INFLV, YSP and UNCf. 'Top factors all' includes the above as well as LIQNT. For the BMA-SDF, we report results for a range prior Sharpe ratio values that are set as 20%, 40%, 60% and 80% of the ex-post maximum Sharpe ratio of the 64 bond portfolios and traded factors. In-sample (IS) test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity and the 14 traded bond factors ($N = 64$). Out-of-sample (OS) test assets include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the 17 Fama French industry portfolios ($N = 77$). In the cross-sectional OS tests, models are first estimated using the baseline IS test assets and then used to price (with no additional parameter estimation) the OS assets. All data is standardized, that is, pricing errors are in Sharpe ratio units.

the BMA-SDF (with a prior Sharpe ratio at 80%) with and without stock-related factors generates an average R^2_{OLS} of 0.33 and 0.30, respectively, with a standard deviation of only 0.12. Similarly, the BMA-SDF achieves superior average performance with a lower standard error across the other asset pricing evaluation metrics. Second, as we ease the shrinkage constraint, the inclusion of equity-related factors improves the asset pricing performance of the BMA-SDF across all metrics considered in Table 4. For an 80% Sharpe ratio prior, the average R^2_{OLS} (R^2_{GLS}) of the BMA-SDF increases from 0.299 (0.202) to 0.330 (0.229). At the same time, the average RMSE (MAPE) decreases from 0.106 (0.082) to 0.103 (0.080), respectively. Similarly, the R^2_{OLS} (R^2_{GLS}) of the top factor model increases from 0.040 (0.121) to 0.193 (0.234) when including the equity factors.

To zoom in on the pricing of the credit risk component of corporate bond returns, we repeat our in- and out-of-sample cross-sectional exercises using duration-adjusted returns as advocated by van Binsbergen, Nozawa, and Schwert (2023). Results are reported in Table 5. Clearly, the BMA-SDFs, with or without equity factors, strongly outperform the benchmark models in- and out-of-sample for any value of the prior Sharpe ratio. Furthermore, the BMA-SDF, in particular when constructed to include both equity and bond factors, strongly outperforms the most likely factor models both in- and out-of-sample. Furthermore, as shown in Table A.4 of Appendix C, the superior cross-sectional OSS performance of the BMA-SDF is very stable across different sets of duration-adjusted returns, and the best performance is achievable using *both* equity and bond factors for its construction.

As an additional robustness check, we repeat our analysis using the publicly available WRDS TRACE database spanning the sample period 2002:09 to 2021:09.[30] Despite loosing more than 16 years of monthly observations compared to our baseline data set, results are quite stable both in terms of most likely components to be included in the SDF and in terms of superior asset pricing performance of the BMA-SDF relative to benchmark models (see Figure IA.2 and Table IA.I of the Internet Appendix).

[30]A detailed description of the data construction is reported in Internet Appendix IA.1.

4 Conclusion

Using a flexible and powerful Bayesian method to study linear factor models, we parse the bond factor zoo. We take a comprehensive view of linear stochastic discount factors (SDFs) to price an extensive cross-section of corporate bond portfolios, using information from the bond as well as the equity market literature.

We find that the majority of tradable factors designed to price corporate bonds are unlikely sources of priced risk, and that only one factor, capturing the post-earnings announcement drift in corporate bonds, which has not been utilized in prior asset pricing models, should be included in any stochastic discount factor (SDF) with very high probability. Furthermore, we find that nontradable factors capturing inflation volatility risk (Kang and Pflueger (2015)) and the term structure yield spread (Koijen, Lustig, and Van Nieuwerburgh (2017)), as well as the return on a broad based bond market index, are likely components of the SDF. Moreover, our results imply that the low dimensional models suggested in the previous literature to price the cross-section of corporate bond returns have very weak support in the data, are misspecified with high probability, and the prices of risk of several factors are only weakly identified in the cross-section of corporate bond returns, hence invalidating canonical inference (but not, importantly, the estimation method we use).

Including factors from the equity factor zoo increases the explanatory power for the cross-section of corporate bond returns, especially when using duration-adjusted test assets. Allowing the SDF to load on equity factors, the ex post maximum Sharpe ratio achievable in the data is no more unrealistically large relative to the posterior distribution of the SDF-implied Sharpe ratio. At the same time, the posterior mean of the number of factors in the SDF is large with 22 to 23 factors, and sparse models have extremely small posterior probabilities, implying that any low-dimensional factor model is extremely likely to be misspecified.

A Bayesian model averaging-stochastic discount factor (BMA-SDF) including bond *and* equity factors prices bonds better than all existing models, both in- and out-of-sample, and has a clear business cycle pattern, increasing during expansions and peaking right before recessions and around the time of financial market crashes. Hence, our approach yields an estimated SDF that behaves as one would expect from the intertemporal marginal rate of substitution of an agent that prices asset returns and is exposed to the risks arising from general economic conditions.

References

Andreani, M., D. Palhares, and S. Richardson (2023): "Computing Corporate Bond Returns: A Word (or Two) of Caution," *Review of Accounting Studies*, forthcoming.

Asness, C., and A. Frazzini (2013): "The Devil in HML's Details," *Journal of Portfolio Management*, 39(4), 49–68.

Asness, C. S., A. Frazzini, and L. H. Pedersen (2019): "Quality Minus Junk," *Review of Accounting Studies*, 24(1), 34–112.

Avramov, D., S. Cheng, L. Metzker, and S. Voigt (2023): "Integrating Factor Models," *The Journal of Finance*, 78(3), 1593–1646.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov (2022): "The Distress Anomaly is Deeper than you Think: Evidence from Stocks and Bonds," *Review of Finance*, 26(2), 355–405.

Bai, J., T. G. Bali, and Q. Wen (2019): "RETRACTED: Common Risk Factors in the Cross-Section of Corporate Bond Returns," *Journal of Financial Economics*, 131(3), 619–642.

Bali, T. G., A. Subrahmanyam, and Q. Wen (2017): "Return-Based Factors for Corporate Bonds," Working Paper.

——— (2021): "Long-Term Reversals in the Corporate Bond Market," *Journal of Financial Economics*, 139(2), 656–677.

Bao, J., J. Pan, and J. Wang (2011): "The Illiquidity of Corporate Bonds," *The Journal of Finance*, 66, 911–946.

Barillas, F., and J. Shanken (2016): "Which Alpha?," *The Review of Financial Studies*, 30(4), 1316–1338.

——— (2018): "Comparing Asset Pricing Models," *The Journal of Finance*, 73(2), 715–754.

Bartram, S. M., M. Grinblatt, and Y. Nozawa (2020): "Book-to-Market, Mispricing, and the Cross-Section of Corporate Bond Returns," Discussion paper, National Bureau of Economic Research.

Bauwens, L., M. Lubrano, and J.-F. Richard (1999): *Bayesian Inference in Dynamic Econometric Models*. Oxford University Press, Oxford.

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *Review of Economic Studies*, 81, 608–650.

Bessembinder, H., K. M. Kahle, W. F. Maxwell, and D. Xu (2008): "Measuring abnormal bond performance," *The Review of Financial Studies*, 22(10), 4219–4258.

Bhamra, H. S., A. J. Fisher, and L.-A. Kuehn (2011): "Monetary policy and corporate default," *Journal of Monetary Economics*, 58(5), 480–494.

Bhamra, H. S., L.-A. Kuehn, and I. A. Strebulaev (2010): "The Levered Equity Risk Premium and Credit Spreads: A Unified Framework," *The Review of Financial Studies*, 23(2), 645–703.

Blume, M. E., and D. B. Keim (1987): "Lower-Grade Bonds: Their Risks and Returns," *Financial Analysts Journal*, 43(4), 26–66.

Boyarchenko, N., L. Elias, and P. Mueller (2023): "Corporate Credit Provision," Working Paper, Warwick Business School.

Bryzgalova, S., J. Huang, and C. Julliard (2022): "Bayesian Fama-MacBeth," Working paper, London School of Economics.

Bryzgalova, S., J. Huang, and C. Julliard (2023): "Bayesian Solutions for the Factor Zoo: We Just Ran Two Quadrillion Models," *Journal of Finance*, 78(1), 487–557.

Campbell, J. Y., and G. B. Taksler (2003): "Equity Volatility and Corporate Bond Yields," *Journal of Finance*, 58, 2321 – 2349.

Carhart, M. M. (1997): "On Persistence in Mutual Fund Performance," *The Journal of Finance*, 52, 57–82.

Ceballos, L. (2022): "Inflation Volatility Risk and the Cross-Section of Corporate Bond Returns," Working Paper, University of San Diego.

Chan, L. K., N. Jegadeesh, and J. Lakonishok (1996): "Momentum Strategies," *Journal of Finance*, 51(5), 1681–1713.

Chen, H., R. Cui, Z. He, and K. Milbradt (2018): "Quantifying Liquidity and Default Risks of Corporate Bonds over the Business Cycle," *Review of Financial Studies*, 31(3), 852–897.

Chib, S., X. Zeng, and L. Zhao (2020): "On Comparing Asset Pricing Models," *The Journal of Finance*, 75(1), 551–577.

Chordia, T., A. Goyal, Y. Nozawa, A. Subrahmanyam, and Q. Tong (2017): "Are Capital Market Anomalies Common to Equity and Corporate Bond Markets? An Empirical Investigation," *Journal of Financial and Quantitative Analysis*, 52(4), 1301–1342.

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

Cochrane, J. H. (2005): *Asset Pricing*, vol. 1. Princeton University Press Princeton, NJ.

Dang, T. D., F. Hollstein, and M. Prokopczuk (2023): "Which Factors for Corporate Bond Returns?," *The Review of Asset Pricing Studies*, forthcoming.

Daniel, K., D. Hirshleifer, and L. Sun (2020): "Short- and Long-Horizon Behavioral Factors," *The Review of Financial Studies*, 33, 1673–1736.

Daniel, K., L. Mota, S. Rottke, and T. Santos (2020): "The Cross-Section of Risk and Returns," *The Review of Financial Studies*, 33(5), 1927–1979.

Della Vigna, S., and J. M. Pollet (2009): "Investor Inattention and Friday Earnings Announcements," *Journal of Finance*, 64(2), 709–749.

Dello-Preite, M., R. Uppal, P. Zaffaroni, and I. Zviadadze (2023): "What is Missing in Asset-Pricing Factor Models?," Working paper, EDHEC Business School.

Dick-Nielsen, J., P. Feldhütter, L. H. Pedersen, and C. Stolborg (2023): "Corporate Bond Factors: Replication Failures and a New Framework," Working Paper.

Dickerson, A., P. Mueller, and C. Robotti (2023): "Priced Risk in Corporate Bonds," *Journal of Financial Economics*, forthcoming.

Dickerson, A., C. Robotti, and G. Rossetti (2023): "Noisy Prices and Return-based Anomalies in Corporate Bonds," Working Paper, Warwick Business School.

Driessen, J. (2005): "Is Default Event Risk Priced in Corporate Bonds?," *Review of Financial Studies*, 18, 165–195.

Elkamhi, R., C. Jo, and Y. Nozawa (2023): "A One-Factor Model of Corporate Bond Premia," *forthcoming, Management Science*.

Elton, E. J., M. J. Gruber, D. Agrawal, and C. Mann (2001): "Explaining the Rate Spread on Corporate Bonds," *Journal of Finance*, 56, 247–277.

Elton, E. J., M. J. Gruber, D. Agrawal, and C. Mann (2004): "Factors affecting the valuation of corporate bonds," *Journal of Banking & Finance*, 28(11), 2747–2767.

Elton, E. J., M. J. Gruber, and C. R. Blake (1995): "Fundamental Economic Variables, Expected Returns, and Bond Fund Performance," *Journal of Finance*, 50(4), 1229–56.

Eom, Y. H., J. Helwege, and J.-z. Huang (2004): "Structural Models of Corporate Bond Pricing: An Empirical Analysis," *Review of Financial Studies*, 17, 499–544.

Fama, E. F., and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47(2), 427–465.

——— (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

——— (2015): "A Five-Factor Asset Pricing Model," *Journal of Financial Economics*, 116(1), 1–22.

Fama, E. F., and J. MacBeth (1973): "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81, 607–636.

Fang, X., Y. Liu, and N. Roussanov (2022): "Getting to the Core: Inflation Risks Within and Across Asset Classes," Discussion paper, National Bureau of Economic Research.

Favilukis, J., X. Lin, and X. Zhao (2020): "The Elephant in the Room: The Impact of Labor Obligations on Credit Markets," *American Economic Review*, 110(6), 1673–1712.

Feldhütter, P., and S. M. Schaefer (2018): "The Myth of the Credit Spread Puzzle," *Review of Financial Studies*, 31(8), 2897–2942.

Feng, G., S. Giglio, and D. Xiu (2020): "Taming the Factor Zoo: A Test of New Factors," *Journal of Finance*, 75(3), 1327–1370.

Fisher, L. (1959): "Determinants of Risk Premiums on Corporate Bonds," *Journal of Political Economy*, 67.

Frazzini, A., and L. H. Pedersen (2014): "Betting Against Beta," *Journal of Financial Economics*, 111(1), 1–25.

Ghosh, A., C. Julliard, and A. P. Taylor (2016): "What Is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models," *The Review of Financial Studies*, 30(2), 442–504.

Giesecke, K., F. A. Longstaff, S. Schaefer, and I. Strebulaev (2011): "Corporate bond default risk: A 150-year perspective," *Journal of Financial Economics*, 102(2), 233–250.

Giglio, S., and D. Xiu (2021): "Asset Pricing with Omitted Factors," *Journal of Political Economy*, 129(7), 1947–1990.

Gomes, J. F., and L. Schmid (2021): "Equilibrium Asset Pricing with Leverage and Default," *Journal of Finance*, 76(2), 977–1018.

Gospodinov, N., R. Kan, and C. Robotti (2014): "Misspecification-Robust Inference in Linear Asset-Pricing Models with Irrelevant Risk Factors," *The Review of Financial Studies*, 27(7), 2139–2170.

——— (2019): "Too Good to Be True? Fallacies in Evaluating Risk Factor Models," *Journal of Financial Economics*, 132(2), 451–471.

Gospodinov, N., and C. Robotti (2021): "Common Pricing Across Asset Classes: Empirical Evidence Revisited," *Journal of Financial Economics*, 140, 292–324.

Hansen, L. P. (1982): "Large Sample Properties of Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Harvey, C. R. (2017): "Presidential Address: The Scientific Outlook in Financial Economics," *The Journal of Finance*, 72(4), 1399–1440.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "...and the Cross-Section of Expected Returns," *The Review of Financial Studies*, 29(1), 5–68.

He, Z., B. Kelly, and A. Manela (2017): "Intermediary Asset Pricing: New Evidence From Many Asset Classes," *Journal of Financial Economics*, 126(1), 1–35.

Hirshleifer, D., S. S. Lim, and S. H. Teoh (2011): "Limited Investor Attention and Stock Market Misreactions to Accounting Information," *Review of Asset Pricing Studies*, 1(1), 35–73.

Hirshleifer, D., and S. H. Teoh (2003): "Limited Attention, Information Disclosure, and Financial Reporting," *Journal of Accounting and Economics*, 36(1-3), 337–386.

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999): "Bayesian Model Averaging: A Tutorial," *Statistical Science*, 14(4), 382–401.

Hou, K., H. Mo, C. Xue, and L. Zhang (2021): "An Augmented q-Factor Model with Expected Growth," *Review of Finance*, 25(1), 1–41.

Hou, K., C. Xue, and L. Zhang (2015): "Digesting Anomalies: An Investment Approach," *The Review of Financial Studies*, 28(3), 650–705.

Houweling, P., and J. Van Zundert (2017): "Factor Investing in the Corporate Bond Market," *Financial Analysts Journal*, 73(2), 100–115.

Jegadeesh, N., and S. Titman (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *Journal of Finance*, 48(1), 65–91.

——— (2001): "Profitability of Momentum Strategies: An Evaluation of Alternative Explanations," *Journal of Finance*, 56(2), 699–720.

Jostova, G., S. Nikolova, A. Philipov, and C. W. Stahel (2013): "Momentum in Corporate Bond Returns," *The Review of Financial Studies*, 26(7), 1649–1693.

Kan, R., and C. Zhang (1999a): "GMM Tests of Stochastic Discount Factor Models with Useless Factors," *Journal of Financial Economics*, 54(1), 103–127.

——— (1999b): "Two-Pass Tests of Asset Pricing Models with Useless Factors," *The Journal of Finance*, 54(1), 203–235.

Kang, J., and C. E. Pflueger (2015): "Inflation Risk in Corporate Bonds," *Journal of Finance*, 70(1), 115–162.

Kelly, B., D. Palhares, and S. Pruitt (2023): "Modeling Corporate Bond Returns," *Journal of Finance*, 78(4), 1967–2008.

Kelly, B. T., and S. Pruitt (2022): "Reconciling TRACE Bond Returns," Working Paper.

Khan, A., and J. K. Thomas (2013): "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity," *Journal of Political Economy*, 121(6), 1055–1107.

Kleibergen, F. (2009): "Tests of Risk Premia in Linear Factor Models," *Journal of Econometrics*, 149(2), 149–173.

Kleibergen, F., and Z. Zhan (2020): "Robust Inference for Consumption-Based Asset Pricing," *The Journal of Finance*, 75(1), 507–550.

Koijen, R. S., H. Lustig, and S. Van Nieuwerburgh (2017): "The Cross-Section and Time Series of Stock and Bond Returns," *Journal of Monetary Economics*, 88, 50–69.

Kozak, S., S. Nagel, and S. Santosh (2020): "Shrinking the Cross-Section," *Journal of Financial Economics*, 135, 271–292.

Lewellen, J., S. Nagel, and J. Shanken (2010): "A Skeptical Appraisal of Asset Pricing Tests," *Journal of Financial Economics*, 96, 175–194.

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

Lintner, J. (1965): "Security Prices, Risk, and Maximal Gains from Diversification," *The Journal of Finance*, 20(4), 587–615.

McCullough, J. R. (1830): *The Principles of Political Economy: With a Sketch of the Rise and Progress of the Science (2nd ed.)*. Edinburgh, London, and Dublin.

Mishkin, F. S., and E. N. White (2002): "U.S. Stock Market Crashes and Their Aftermath: Implications for Monetary Policy," NBER Working Papers 8992, National Bureau of Economic Research, Inc.

Newey, W. K., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, ed. by R. F.Engle, and D. McFadden, vol. 4. Elsevier Press.

Nozawa, Y. (2017): "What Drives the Cross-Section of Credit Spreads?: A Variance Decomposition Approach," *The Journal of Finance*, 72(5), 2045–2072.

Nozawa, Y., Y. Qiu, and Y. Xiong (2023): "Disagreement and Price Drifts in the Corporate Bond Market," Working Paper, University of Toronto.

Palhares, D., and S. Richardson (2020): "Looking Under the Hood of Active Credit Managers," *Financial Analysts Journal*, 76(2), 82–102.

Parker, J. A., and C. Julliard (2003): "Consumption Risk and Cross-Sectional Returns," Working Paper 9538, National Bureau of Economic Research.

Parker, J. A., and C. Julliard (2005): "Consumption Risk and the Cross Section of Expected Returns," *Journal of Political Economy*, 113(1), 185–222.

Pástor, L. (2000): "Portfolio Selection and Asset Pricing Models," *The Journal of Finance*, 55(1), 179–223.

Pástor, L., and R. F. Stambaugh (2000): "Comparing Asset Pricing Models: An Investment Perspective," *Journal of Financial Economics*, 56(3), 335–381.

——— (2003): "Liquidity Risk and Expected Stock Returns," *Journal of Political Economy*, 111, 642–685.

Raftery, A. E., D. Madigan, and J. A. Hoeting (1997): "Bayesian Model Averaging for Linear Regression Models," *Journal of the American Statistical Association*, 92(437), 179–191.

Raftery, A. E., and Y. Zheng (2003): "Discussion: Performance of Bayesian Model Averaging," *Journal of the American Statistical Association*, 98, 931–938.

Schervish, M. J. (1995): *Theory of Statistics*, Springer Series in Statistics. Springer-Verlag.

Sharpe, W. F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *The Journal of Finance*, 19(3), 425–442.

van Binsbergen, J. H., Y. Nozawa, and M. Schwert (2023): "Duration-Based Valuation of Corporate Bonds," Jacobs Levy Equity Management Center for Quantitative Financial Research Paper.

Appendix

A Posterior sampling

The posterior of the time series parameters follows the the canonical Normal-inverse-Wishart distribution (see, e.g., Bauwens, Lubrano, and Richard 1999) given by

$$\boldsymbol{\mu_Y}|\boldsymbol{\Sigma_Y}, \mathbf{Y} \sim \mathcal{N}\left(\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}},\ \ \boldsymbol{\Sigma_Y}/T\right), \tag{A.7}$$

$$\boldsymbol{\Sigma_Y}|\mathbf{Y} \sim \mathcal{W}^{\text{-}1}\left(T-1,\ \ \sum_{t=1}^{T}\left(\boldsymbol{Y_t}-\widehat{\boldsymbol{\mu}}_{\boldsymbol{Y}}\right)\left(\boldsymbol{Y_t}-\widehat{\boldsymbol{\mu}}_{\boldsymbol{Y}}\right)^{\top}\right), \tag{A.8}$$

where $\hat{\boldsymbol{\mu}}_{\boldsymbol{Y}} \equiv \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{Y_t}$, $\mathcal{W}^{\text{-}1}$ is the inverse-Wishart distribution, $\boldsymbol{Y} \equiv \{\boldsymbol{Y}_t\}_{t=1}^{T}$, and note that the covariance matrix of factors and test assets, $\boldsymbol{C_f}$, is contained within $\boldsymbol{\Sigma_Y}$.

Define $\boldsymbol{D}$ as a diagonal matrix with elements $c,\ (r(\gamma_1)\psi_1)^{-1}, \dots,\ (r(\gamma_K)\psi_K)^{-1}$. Hence, in matrix notation, the prior for $\boldsymbol{\lambda}$ in equation (3) is $\boldsymbol{\lambda}|\sigma^2, \boldsymbol{\gamma} \sim \mathcal{N}(0, \sigma^2\boldsymbol{D}^{-1})$. It then follows that, given our prior formulations, the posterior distributions of the parameters in the cross-sectional layer $(\boldsymbol{\lambda}, \boldsymbol{\gamma}, \boldsymbol{\omega}, \sigma^2)$, conditional on the draws of $\boldsymbol{\mu_R}$, $\boldsymbol{\Sigma_R}$, and $\boldsymbol{C}$ from the time series layer, are (see Bryzgalova, Huang, and Julliard 2023 for a formal derivation):

$$\boldsymbol{\lambda}|\text{data}, \sigma^2, \boldsymbol{\gamma}, \boldsymbol{\omega} \sim \mathcal{N}\left(\hat{\boldsymbol{\lambda}}, \hat{\sigma}^2(\hat{\boldsymbol{\lambda}})\right), \tag{A.9}$$

$$\frac{p(\gamma_j = 1|\text{data}, \boldsymbol{\lambda}, \boldsymbol{\omega}, \sigma^2, \boldsymbol{\gamma_{-j}})}{p(\gamma_j = 0|\text{data}, \boldsymbol{\lambda}, \boldsymbol{\omega}, \sigma^2, \boldsymbol{\gamma_{-j}})} = \frac{\omega_j}{1-\omega_j}\frac{p(\lambda_j|\gamma_j = 1, \sigma^2)}{p(\lambda_j|\gamma_j = 0, \sigma^2)}, \tag{A.10}$$

$$\omega_j|\text{data}, \boldsymbol{\lambda}, \boldsymbol{\gamma}, \sigma^2 \sim Beta\left(\gamma_j + a_\omega, 1 - \gamma_j + b_\omega\right), \tag{A.11}$$

$$\sigma^2|\text{data}, \boldsymbol{\omega}, \boldsymbol{\lambda}, \boldsymbol{\gamma} \sim \mathcal{IG}\left(\frac{N+K+1}{2}, \frac{(\boldsymbol{\mu_R} - \boldsymbol{C\lambda})^{\top}\boldsymbol{\Sigma_R}^{-1}(\boldsymbol{\mu_R} - \boldsymbol{C\lambda}) + \boldsymbol{\lambda}^{\top}\boldsymbol{D\lambda}}{2}\right), \tag{A.12}$$

where $\hat{\boldsymbol{\lambda}} = (\boldsymbol{C}^{\top}\boldsymbol{\Sigma_R}^{-1}\boldsymbol{C} + \boldsymbol{D})^{-1}\boldsymbol{C}^{\top}\boldsymbol{\Sigma_R}^{-1}\boldsymbol{\mu_R}$, $\hat{\sigma}^2(\hat{\boldsymbol{\lambda}}) = \sigma^2(\boldsymbol{C}^{\top}\boldsymbol{\Sigma_R}^{-1}\boldsymbol{C} + \boldsymbol{D})^{-1}$.

Hence, posterior sampling is achieved with a Gibbs sampler that draws sequentially the time series layer parameters ($\boldsymbol{\mu_R}$, $\boldsymbol{\Sigma_R}$, and $\boldsymbol{C}$) from equations (A.7)-(A.8), and then, conditional on these realizations, draws sequentially from equations (A.9)–(A.12).

B The factor zoo

We present the 49 bond and equity factors used in Table A.1 including a detailed description of their construction, associated reference, and data source.

Table A.1: List of factors for cross-sectional asset pricing. This table presents the list of tradable bond, equity and nontradable factors used in the main paper. For each of the factors, we present their identification index (Factor ID), a description of the factor construction, and the source of the data for downloading and/or constructing the time series.

Factor ID	**Factor name and description**	**Reference**	**Source**
	Panel A: Traded corporate bond factors		
CRF	Credit risk factor. Equally-weighted average return on two 'credit portfolios': CRF_{VaR}, and CRF_{REV}. CRF_{VaR} is the average return difference between the lowest-rating (i.e., highest credit risk) portfolio and the highest-rating (i.e., lowest credit risk) portfolio across the VaR95 portfolios. CRF_{REV} is the average return difference between the lowest-rating portfolio and the highest-rating portfolio across quintiles sorted on bond short-term reversal.	Bai, Bali, and Wen (2019)	Open Source Bond Asset Pricing
CRY	Bond carry factor. Independent sort (5×5) to form 25 portfolios according to ratings and bond credit spreads (CS). For each rating quintile, calculate the weighted average return difference between the highest CS quintile and the lowest CS quintile. CRY is computed as the average long-short portfolio return across all rating quintiles.	Houweling and Van Zundert (2017)	Open Source Bond Asset Pricing
DEF	Bond default risk factor. The difference between the return on the market portfolio of long-term corporate bond returns (the Composite portfolio on the corporate bond module of Ibbotson Associates) and the long-term government bond return.	Fama and French (1992)	Amit Goyal website
DRF	Downside risk factor. Independent sort (5×5) to form 25 portfolios according to ratings and 95% value-at-risk (VaR95). For each rating quintile, calculate the weighted average return difference between the highest VaR5 quintile and the lowest VaR5 quintile. DRF is computed as the average long-short portfolio return across all rating quintiles.	Bai, Bali, and Wen (2019)	Open Source Bond Asset Pricing
DUR	Bond duration factor. Independent sort (5×5) to form 25 portfolios according to ratings and bond duration (DUR^B). For each rating quintile, calculate the weighted average return difference between the highest DUR^B quintile and the lowest DUR^B quintile. DUR is computed as the average long-short portfolio return across all rating quintiles.	Dang, Hollstein, and Prokopczuk (2023)	Open Source Bond Asset Pricing
HMLB	Bond book-to-market factor. Independent sort (2×3) to form 6 portfolios according to bond size and bond book-to-market (BBM), defined as bond principal value scaled by market value. For each size portfolio, calculate the weighted average return difference between the lowest BBM tercile and the highest BBM tercile. HMLB is computed as the average long-short portfolio return across the two size portfolios.	Bartram, Grinblatt, and Nozawa (2020)	Open Source Bond Asset Pricing

LTREVB	Bond long-term reversal factor. Dependent sort ($3 \times 3 \times 3$) to form 27 portfolios according to ratings, maturity, and the 48-13 cumulative previous bond return (LTREV^B). For each rating quintile, the factor is computed as the average return differential between the portfolio with the lowest LTREV^B and the one with the highest LTREV^B within the rating and maturity portfolios. LTREVB is computed as the average long-short portfolio return across the nine rating-maturity terciles.	Bali, Subrahmanyam, and Wen (2021)	Open Source Bond Asset Pricing
MKTB	Corporate Bond Market excess return. Constructed using bond returns in excess of the one-month risk-free rate of return.	Dickerson, Mueller, and Robotti (2023)	Open Source Bond Asset Pricing
MKTBD	Corporate Bond Market duration adjusted return. Constructed using bond returns in excess of their duration-matched U.S. Treasury bond rate of return.	van Binsbergen, Nozawa, and Schwert (2023)	Open Source Bond Asset Pricing
MOMB	Bond momentum factor. Independent sort (5×5) to form 25 portfolios according to ratings and the 12-2 cumulative previous bond return (MOM). For each rating quintile, calculate the weighted average return difference between the highest MOM quintile and the lowest MOM quintile. MOMB is computed as the average long-short portfolio return across all rating quintiles.	Bali, Subrahmanyam, and Wen (2017)	Open Source Bond Asset Pricing
PEADB	Bond earnings announcement drift factor. Independent sort (2×3) to form 6 portfolios according to market equity and earnings surprises (CAR), computed according to Chan, Jegadeesh, and Lakonishok (1996). For each firm size portfolio, calculate the weighted average return difference between the highest CAR terciles and the lowest CAR tercile. PEADB is computed as the average long-short portfolio return across the two firm size portfolios.	Nozawa, Qiu, and Xiong (2023)	Open Source Bond Asset Pricing
STREVB	Bond short-term reversal factor. Independent sort (5×5) to form 25 portfolios according to ratings and the prior month's bond return (REV). For each rating quintile, calculate the weighted average return difference between the lowest REV quintile and the highest REV quintile. STREVB is computed as the average long-short portfolio return across all rating quintiles.	Bali, Subrahmanyam, and Wen (2021)	Open Source Bond Asset Pricing
TERM	Bond term structure risk factor. The difference between the monthly long-term government bond return and the one-month T-Bill rate of return.	Fama and French (1992)	Amit Goyal website
VAL	Bond value factor. Independent sort (2×3) to form 6 portfolios according to bond size and bond value (VAL^B). VAL^B is computed via cross-sectional regressions of credit spreads on ratings, maturity, and the 3-month change in credit spread. The percentage difference between the actual credit spread and the fitted ('fair') credit spread for each bond is the VAL^B characteristic. For each size portfolio, calculate the weighted average return difference between the highest VAL^B tercile and the lowest VAL^B tercile. VAL is computed as the average long-short portfolio return across the two size portfolios.	Houweling and Van Zundert (2017)	Open Source Bond Asset Pricing
	Panel B: Nontraded corporate bond and equity factors		
CPTL	Intermediary capital nontraded risk factor. Constructed using AR(1) innovations to the market-based capital ratio of primary dealers, scaled by the lagged capital ratio.	He, Kelly, and Manela (2017)	Zhiguo He website
EPU	Economic Policy Uncertainty. First difference in the economic policy uncertainty index.	Dang, Hollstein, and Prokopczuk (2023)	FRED
EPUT	Economic Tax Policy Uncertainty. First difference in the economic tax policy uncertainty index.	Dang, Hollstein, and Prokopczuk (2023)	FRED
INFLC	Shocks to core inflation. Unexpected core inflation component captured by an ARMA(1,1) model. Monthly core inflation is calculated as the percentage change in the seasonally adjusted Consumer Price Index for All Urban Consumers: All Items Less Food and Energy which is lagged by one-month to account for the inflation data release lag.	Fang, Liu, and Roussanov (2022)	FRED

INFLV	Inflation volatility. Computed as the 6-month volatility of the unexpected inflation component captured by an ARMA(1,1) model. Monthly inflation is calculated as the percentage change in the seasonally adjusted Consumer Price Index for All Urban Consumers (CPI) which is lagged by one-month to account for the inflation data release lag.	Kang and Pflueger (2015) and Ceballos (2022)	FRED
LVL	Level term structure factor. Constructed as the first principal component of the one- through 30-year CRSP Fixed Term Indices U.S. Treasury Bond yields.	Koijen, Lustig, and Van Nieuwerburgh (2017)	CRSP Indices
LIQNT	Liquidity factor, computed as the average of individual-stock measures estimated with daily data (residual predictability, controlling for the market factor)	Pástor and Stambaugh (2003)	Robert Stambaugh website
UNC	First difference in the Macroeconomic uncertainty index, which is lagged by one-month to align the forecast to the returns observed in month t.	Koijen, Lustig, and Van Nieuwerburgh (2017)	Sydney Ludvigson website
UNCf	First difference in the Financial economic uncertainty index, which is lagged by one-month to align the forecast to the returns observed in month t.	Koijen, Lustig, and Van Nieuwerburgh (2017)	Sydney Ludvigson website
UNCr	First difference in the Real economic uncertainty index, which is lagged by one-month to align the forecast to the returns observed in month t.	Koijen, Lustig, and Van Nieuwerburgh (2017)	Sydney Ludvigson website
VIX	First difference in the CBOE VIX.	Chung, Wang, and Wu (2019)	FRED
YSP	Slope term structure factor. Constructed as the difference in the five and one-year U.S. Treasury Bond yields.	Koijen, Lustig, and Van Nieuwerburgh (2017)	CRSP Indices
Panel C: Traded equity factors			
BAB	Betting-against-beta factor, constructed as a portfolio that holds low-beta assets, leveraged to a beta of 1, and that shorts high-beta assets, de-leveraged to a beta of 1	Frazzini and Pedersen (2014)	AQR data library
CMA	Investment factor, constructed as a long-short portfolio of stocks sorted by their investment activity	Fama and French (2015)	Ken French website
CMAs	CMA with a hedged unpriced component	Daniel, Mota, Rottke, and Santos (2020)	Kent Daniel website
CPTLT	The value-weighted equity return for the New York Fed's primary dealer sector not including new equity issuance	He, Kelly, and Manela (2017)	Zhiguo He website
FIN	Long-term behavioral factor, predominantly capturing the impact of share issuance and correction	Daniel, Hirshleifer, and Sun (2020)	Kent Daniel website
HML	Value factor, constructed as a long-short portfolio of stocks sorted by their book-to-market ratio	Fama and French (1992)	Ken French website
HML_DEV	A version of the HML factor that relies on the current price level to sort the stocks into long and short legs	Asness and Frazzini (2013)	AQR data library
HMLs	HML with a hedged unpriced component	Daniel, Mota, Rottke, and Santos (2020)	Kent Daniel website
LIQ	Liquidity factor, constructed as a long-short portfolio of stocks sorted by their exposure to LIQ_NT	Pástor and Stambaugh (2003)	Robert Stambaugh website
LTREV	Long-term reversal factor, constructed as a long-short portfolio of stocks sorted by their cumulative return accrued in the previous 60-13 months	Jegadeesh and Titman (2001)	Ken French website
MKTS	Market excess return	Sharpe (1964) and Lintner (1965)	Ken French website
MKTSs	Market factor with a hedged unpriced component	Daniel, Mota, Rottke, and Santos (2020)	Kent Daniel website
MOMS	Momentum factor, constructed as a long-short portfolio of stocks sorted by their 12-2 cumulative previous return	Carhart (1997), Jegadeesh and Titman (1993)	Ken French website
PEAD	Short-term behavioral factor, reflecting post-earnings announcement drift	Daniel, Hirshleifer, and Sun (2020)	Kent Daniel website
QMJ	Quality-minus-junk factor, constructed as a long-short portfolio of stocks sorted by the combination of their safety, profitability, growth, and the quality of management practices	Asness, Frazzini, and Pedersen (2019)	AQR data library
RMW	Profitability factor, constructed as a long-short portfolio of stocks sorted by their profitability	Fama and French (2015)	Ken French website
RMWs	RMW with a hedged unpriced component	Daniel, Mota, Rottke, and Santos (2020)	Kent Daniel website
R_EG	Expected growth factor, constructed as a long-short portfolio of stocks sorted by their forecasted growth rates	Hou, Mo, Xue, and Zhang (2021)	Lu Zhang website

R_IA	Investment factor, constructed as a long-short portfolio of stocks sorted by their investment-to-capital	Hou, Xue, and Zhang (2015)	Lu Zhang website
R_ROE	Profitability factor, constructed as a long-short portfolio of stocks sorted by their return on equity	Hou, Xue, and Zhang (2015)	Lu Zhang website
SMB	Size factor, constructed as a long-short portfolio of stocks sorted by their market cap	Fama and French (1992)	Ken French website
SMBs	SMB with a hedged unpriced component	Daniel, Mota, Rottke, and Santos (2020)	Kent Daniel website
STREV	Short-term reversal factor, constructed as a long-short portfolio of stocks sorted by their previous month return	Jegadeesh and Titman (1993)	Ken French website

C Results with duration-adjusted bond returns

We replicate the main results in in Section 3 using bond duration-adjusted returns. Duration-adjusted returns are computed for each bond at each time t such that the resultant return is in excess of a portfolio of duration-matched U.S. Treasury Bond returns. The total return for corporate bond i in month t is,

$$R_{i,t} = \frac{B_{i,t} + AI_{i,t} + Coupon_{i,j,t}}{B_{i,t-1} + AI_{i,t-1}} - 1,$$

where $B_{i,t}$ is the clean price of bond i in month t, $AI_{i,t}$ is the accrued interest, and $Coupon_{i,t}$ is the coupon payment, if any.

The bond credit excess return ('duration adjusted return') is the total bond return minus a hedging portfolio of U.S Treasury Bonds that have the same duration as the bond in month t. The duration-adjusted return isolates the portion of a bonds performance that is attributed solely to the credit risk of each bond. The duration-adjusted return is defined as,

$$r^e_{i,t} = R_{i,t} - r^{Dur}_{i,t},$$

where $r_{i,t}$ is the total return of bond i in month t and $r^{Dur}_{i,t}$ is the duration-matched portfolio of U.S Treasury bonds for bond i in month t. We also use $r^e_{i,t}$ to compute the traded bond factor returns.

Results obtained with duration-adjusted returns are reported in Tables A.2–A.4 and Figures A.1–A.7.

Table A.2: Posterior factor probabilities and risk prices – bond specific factor zoo – duration-adjusted returns

	Factor prob., $\mathbb{E}[\gamma_j \mid \text{data}]$				Price of Risk, $\mathbb{E}[\lambda_j \mid \text{data}]$			
	Total Prior SR				Total Prior SR			
Factors:	20%	40%	60%	80%	20%	40%	60%	80%
PEADB	0.589	0.737	0.806	0.797	0.098	0.356	0.602	0.731
MKTB	0.535	0.566	0.575	0.550	0.055	0.159	0.269	0.371
YSP	0.499	0.505	0.519	0.570	0.002	0.011	0.030	0.095
INFLV	0.498	0.508	0.519	0.548	0.005	0.021	0.054	0.139
UNCf	0.503	0.511	0.520	0.515	−0.013	−0.047	−0.106	−0.207
LVL	0.496	0.494	0.496	0.500	0.000	−0.002	−0.005	−0.015
INFLC	0.497	0.495	0.494	0.495	0.000	0.000	0.000	0.001
EPU	0.495	0.490	0.487	0.474	0.001	0.006	0.019	0.050
UNCr	0.491	0.490	0.485	0.477	0.000	0.001	0.006	0.021
EPUT	0.492	0.489	0.487	0.469	0.000	0.000	−0.006	−0.02
UNC	0.493	0.489	0.474	0.441	−0.003	−0.008	−0.011	−0.004
TERM	0.503	0.499	0.460	0.397	0.047	0.117	0.150	0.145
VAL	0.488	0.481	0.464	0.415	0.040	0.101	0.173	0.237
CPTL	0.489	0.480	0.458	0.411	−0.003	−0.014	−0.024	−0.033
CRY	0.488	0.479	0.458	0.410	0.040	0.100	0.165	0.223
VIX	0.488	0.470	0.446	0.393	−0.009	−0.025	−0.046	−0.059
DEF	0.478	0.468	0.451	0.400	−0.022	−0.088	−0.164	−0.217
MKTBD	0.478	0.451	0.425	0.384	0.023	0.027	−0.001	−0.074
DRF	0.480	0.453	0.425	0.373	0.020	0.017	−0.015	−0.074
STREVB	0.487	0.462	0.417	0.343	0.012	0.034	0.055	0.067
HMLB	0.495	0.466	0.410	0.328	0.041	0.077	0.081	0.067
CRF	0.471	0.434	0.394	0.331	0.022	0.065	0.115	0.143
DUR	0.475	0.440	0.390	0.307	−0.014	−0.045	−0.064	−0.060
LTREVB	0.481	0.439	0.374	0.286	0.014	0.017	0.006	−0.003
MOMB	0.481	0.431	0.371	0.288	−0.020	−0.021	0.000	0.018

Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, and posterior mean of (annualized) risk prices, $\mathbb{E}[\lambda_j|\text{data}]$, of the 25 bond specific factors described in Appendix B. The prior for each factor inclusion is a Beta(1, 1), yielding a prior expectation for γ_j of 50%. Test assets are the returns, computed in excess of the duration-matched U.S. Treasury Bond rate of return, of 50 bond portfolios sorted on credit spreads, size, rating and maturity, plus the 14 traded bond factors ($N = 64$). Results are tabulated for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, with values set to 20%, 40%, 60% and 80% of the ex-post maximum Sharpe ratio of the test assets. Sample: 1986:01 to 2021:09 ($T = 429$).

Table A.3: Posterior factor probabilities and risk prices – bond and stock factor zoo – duration-adjusted returns

	Factor prob., $\mathbb{E}[\gamma_j\|\text{data}]$				Price of Risk, $\mathbb{E}[\lambda_j\|\text{data}]$			
	Total Prior SR				Total Prior SR			
Factors:	20%	40%	60%	80%	20%	40%	60%	80%
PEADB	0.571	0.712	0.800	0.822	0.075	0.291	0.546	0.746
MKTB	0.522	0.561	0.568	0.557	0.042	0.132	0.232	0.343
YSP	0.501	0.502	0.515	0.552	0.002	0.008	0.023	0.068
INFLV	0.498	0.504	0.513	0.540	0.004	0.016	0.041	0.103
UNCf	0.501	0.507	0.517	0.514	−0.010	−0.037	−0.084	−0.168
LIQNT	0.495	0.500	0.508	0.530	−0.002	−0.011	−0.032	−0.091
RMWs	0.497	0.500	0.499	0.498	0.000	0.001	0.002	0.006
CMAs	0.498	0.500	0.498	0.494	0.000	0.000	0.000	−0.001
LVL	0.494	0.494	0.496	0.504	0.000	−0.001	−0.004	−0.011
SMBs	0.496	0.495	0.496	0.499	0.001	0.004	0.010	0.027
R_IA	0.498	0.495	0.494	0.494	−0.001	−0.003	−0.007	−0.019
LTREV	0.494	0.495	0.497	0.494	0.002	0.010	0.029	0.073
MKTSs	0.499	0.498	0.492	0.491	−0.001	−0.003	−0.009	−0.026
SMB	0.494	0.492	0.497	0.494	0.003	0.015	0.044	0.100
RMW	0.495	0.494	0.494	0.494	0.003	0.012	0.031	0.079
INFLC	0.494	0.494	0.494	0.493	0.000	0.000	0.000	0.001
CMA	0.494	0.494	0.494	0.491	0.000	−0.002	−0.005	−0.014
HMLs	0.491	0.493	0.494	0.490	−0.001	−0.003	−0.009	−0.027
HML	0.495	0.495	0.491	0.483	0.000	−0.001	−0.001	−0.004
EPU	0.496	0.492	0.490	0.478	0.001	0.005	0.014	0.037
EPUT	0.493	0.494	0.490	0.476	0.000	0.000	−0.004	−0.018
UNCr	0.494	0.492	0.487	0.481	0.000	0.001	0.004	0.013
FIN	0.493	0.490	0.490	0.480	0.001	0.005	0.014	0.034
BAB	0.494	0.492	0.490	0.476	0.002	0.008	0.017	0.030
STREV	0.493	0.491	0.486	0.475	0.003	0.011	0.027	0.062
LIQ	0.496	0.494	0.489	0.467	0.000	−0.002	−0.004	−0.005
PEAD	0.493	0.491	0.483	0.474	0.000	0.002	0.006	0.018
R_EG	0.494	0.493	0.481	0.455	0.000	0.000	−0.003	−0.004
UNC	0.494	0.491	0.478	0.452	−0.002	−0.007	−0.010	−0.008
HMLD	0.493	0.486	0.479	0.454	0.003	0.012	0.028	0.053
TERM	0.504	0.500	0.478	0.428	0.037	0.103	0.150	0.165
CPTL	0.493	0.486	0.471	0.447	−0.002	−0.012	−0.024	−0.044
R_ROE	0.495	0.485	0.473	0.443	−0.003	−0.006	−0.013	−0.026
QMJ	0.490	0.480	0.470	0.450	−0.006	−0.016	−0.04	−0.092
MOMS	0.491	0.484	0.470	0.444	−0.002	−0.006	−0.013	−0.022
MKTS	0.489	0.482	0.469	0.444	0.001	−0.002	−0.010	−0.021
CPTLT	0.493	0.482	0.468	0.440	0.000	−0.004	−0.004	0.005
CRY	0.492	0.481	0.465	0.427	0.033	0.083	0.143	0.213
VAL	0.491	0.479	0.463	0.423	0.032	0.082	0.142	0.199
VIX	0.495	0.479	0.462	0.418	−0.007	−0.019	−0.039	−0.054
DEF	0.476	0.472	0.462	0.428	−0.017	−0.073	−0.148	−0.221
MKTBD	0.482	0.459	0.435	0.401	0.019	0.026	0.004	−0.067
DRF	0.485	0.460	0.435	0.396	0.017	0.019	−0.008	−0.064
HMLB	0.494	0.471	0.427	0.357	0.033	0.068	0.079	0.068
STREVB	0.485	0.467	0.427	0.360	0.010	0.027	0.046	0.056
DUR	0.482	0.451	0.405	0.336	−0.011	−0.040	−0.060	−0.062
CRF	0.475	0.439	0.407	0.350	0.018	0.051	0.097	0.130
LTREVB	0.487	0.451	0.394	0.317	0.012	0.017	0.005	−0.011
MOMB	0.485	0.447	0.393	0.320	−0.017	−0.023	−0.006	0.013

Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, and posterior mean of (annualized) risk prices, $\mathbb{E}[\lambda_j|\text{data}]$, of the 49 bond and equity factors described in AppendixB. The prior for each factor inclusion is a Beta(1, 1), yielding a prior expectation for γ_j of 50%. Test assets are the returns, computed in excess of the duration-matched U.S. Treasury Bond rate of return, of 50 bond portfolios sorted on credit spreads, size, rating and maturity, plus the 14 traded bond factors ($N = 64$). Results are tabulated for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, with values set to 20%, 40%, 60% and 80% of the ex-post maximum Sharpe ratio of the test assets. Sample: 1986:01 to 2021:09 ($T = 429$).

Table A.4: Out-of-sample cross-sectional asset pricing with 127 cross-sections – duration-adjusted

	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}
Panel A: BMA-SDF with 25 bond factors (33.6 mn models)				
prior $SR = 20\%$	0.089	0.065	0.114	0.066
	[0.009]	[0.007]	[0.105]	[0.059]
prior $SR = 40\%$	0.086	0.064	0.128	0.071
	[0.008]	[0.006]	[0.312]	[0.069]
prior $SR = 60\%$	0.088	0.066	0.066	0.056
	[0.010]	[0.009]	[0.488]	[0.075]
prior $SR = 80\%$	0.089	0.067	0.034	0.057
	[0.011]	[0.010]	[0.579]	[0.081]
Panel B: BMA-SDF with 49 bond and stock factors (563 tn models)				
prior $SR = 20\%$	0.089	0.066	0.101	0.063
	[0.010]	[0.007]	[0.082]	[0.056]
prior $SR = 40\%$	0.086	0.063	0.146	0.080
	[0.008]	[0.006]	[0.254]	[0.067]
prior $SR = 60\%$	0.087	0.065	0.112	0.079
	[0.009]	[0.008]	[0.420]	[0.074]
prior $SR = 80\%$	0.087	0.066	0.089	0.092
	[0.011]	[0.010]	[0.524]	[0.081]
Panel C: Benchmark models and most likely factors				
CAPMB	0.095	0.072	−0.025	0.062
	[0.011]	[0.010]	[0.110]	[0.062]
CAPM	0.097	0.073	−0.050	0.039
	[0.011]	[0.010]	[0.082]	[0.060]
FF5	0.101	0.075	−0.159	−0.019
	[0.014]	[0.010]	[0.174]	[0.102]
PEADB	0.095	0.071	−0.045	−0.139
	[0.011]	[0.008]	[0.301]	[0.091]
Top factors bond	0.112	0.094	−0.531	−0.134
	[0.013]	[0.014]	[0.907]	[0.154]
Top factors all	0.128	0.105	−1.079	−0.057
	[0.016]	[0.015]	[1.861]	[0.167]

This table reports the average out-of-sample asset pricing performance and associated standard errors of the BMA-SDF, notable factor models, and factors with a posterior probability greater than 50%. Test assets and the traded bond factors are computed with returns in excess of the duration-matched U.S. Treasury Bond rate of return. The metrics are averaged over 127 possible combinations of the 7 sets of out-of-sample test assets. The out-of-sample (OS) portfolios which form the combinations, include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the Fama French 17 industry portfolios. We use GMM-GLS to estimate factor prices of risk for the CAPMB (bond CAPMB with duration adjusted returns), the equity CAPM, the original FF5 model of Fama and French (1992) and Fama and French (1993), which includes the MKTS, SMB, HML and the DEF and TERM factors, and a single-factor model with PEADB. The 'Top factors bond' includes PEADB, MKTB, INFLV, YSP and UNCf. The 'Top factors all' includes the above as well as LIQNT. For the BMA-SDF, we report results for a range prior Sharpe ratio values that are set as 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. In the cross-sectional OS tests, the models are first estimated using the baseline IS test assets and then used to price (without additional parameters estimation), the OS assets. All the data is standardized, that is, pricing errors are in Sharpe ratio units.

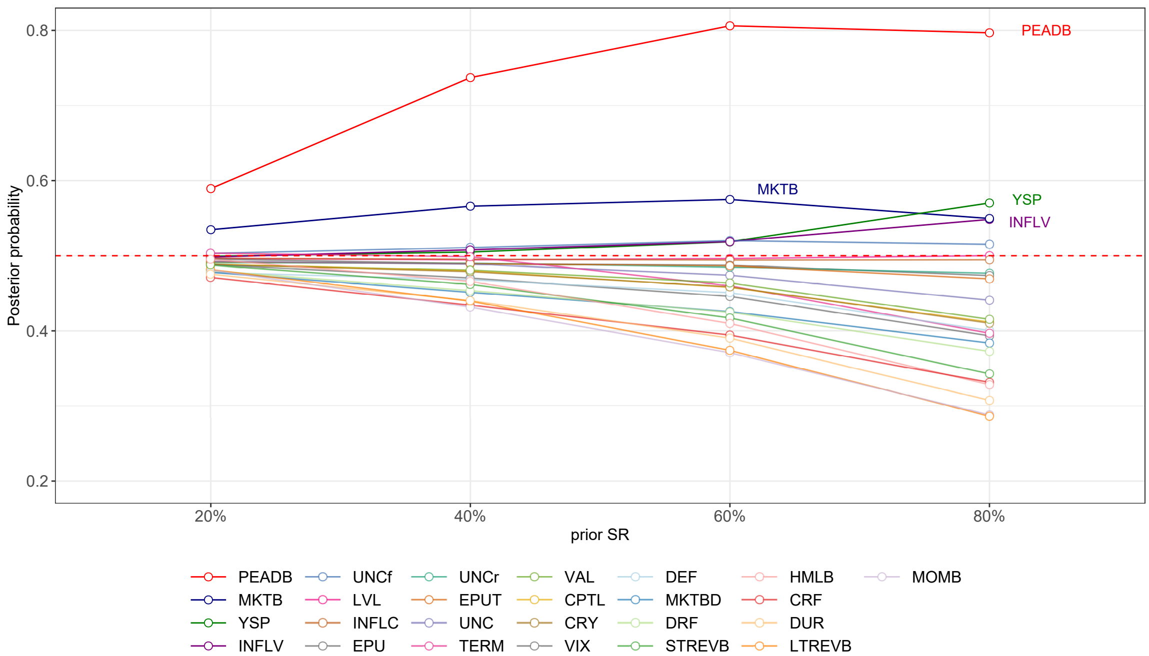


Figure A.1: Posterior factor probabilities – bond factor zoo – duration-adjusted returns.

Posterior probabilities of factors, $\mathbb{E}[\gamma_j|\text{data}]$, of 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors ($N = 64$). Both the traded bond factors and the 50 bond portfolios are constructed in excess of the duration-matched U.S. Treasury Bond rate of return. The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. Posterior probabilities reported for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, set to 20%, 40%, 60% and 80% of the ex-post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

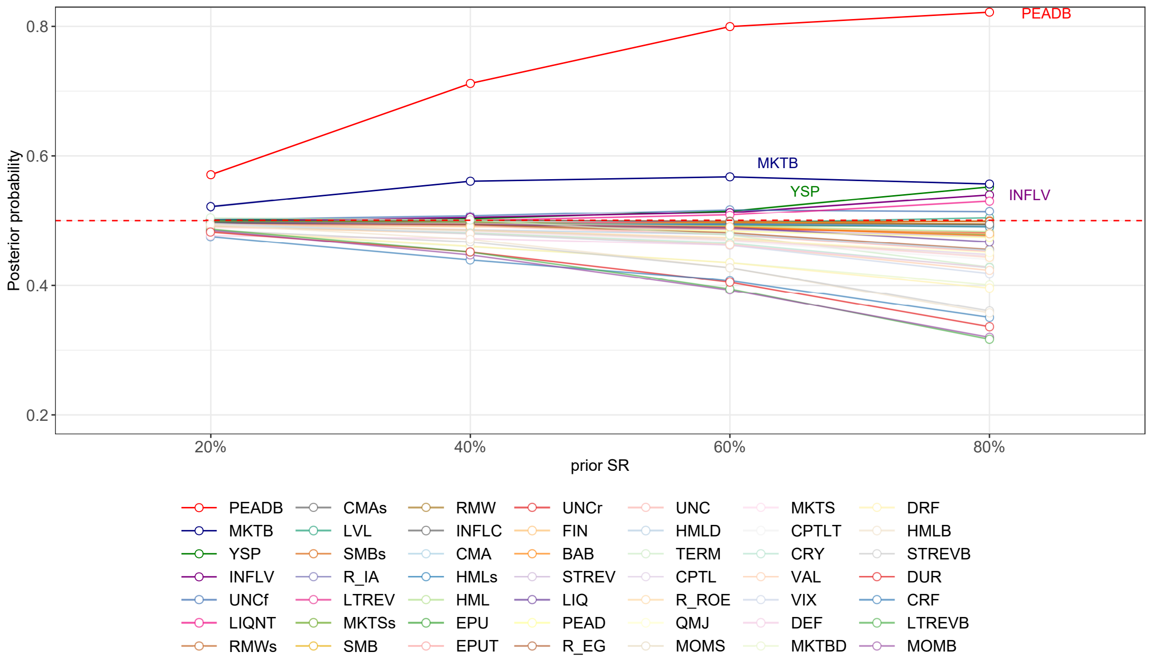


Figure A.2: Posterior factor probabilities – bond and stock factor zoo – duration-adjusted returns.

Posterior probabilities of factors, $\mathbb{E}[\gamma_j|\text{data}]$, of 49 bond and stock factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors ($N = 64$). Both the traded bond factors and the 50 bond portfolios are constructed in excess of the duration-matched U.S. Treasury Bond rate of return. The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. Posterior probabilities for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR^2_{\boldsymbol{f}} \mid \sigma^2]}$, set to 20%, 40%, 60% and 80% of the ex-post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

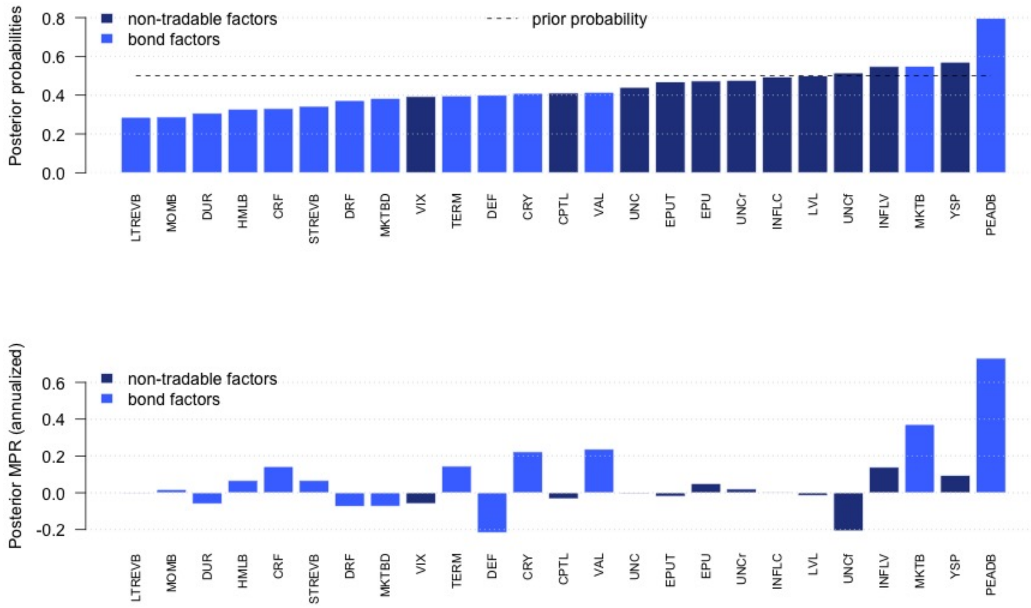


Figure A.3: Posterior factor probabilities and market prices of risk with duration-adjusted returns – bond factor zoo.

Posterior factor probabilities (top panel), $\mathbb{E}[\gamma_j|\text{data}]$, and corresponding posterior market prices of risk (bottom panel), $\mathbb{E}[\lambda_j|\text{data}]$, of 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors, with returns computed in excess of the duration-matched U.S. Treasury Bond rate of return ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

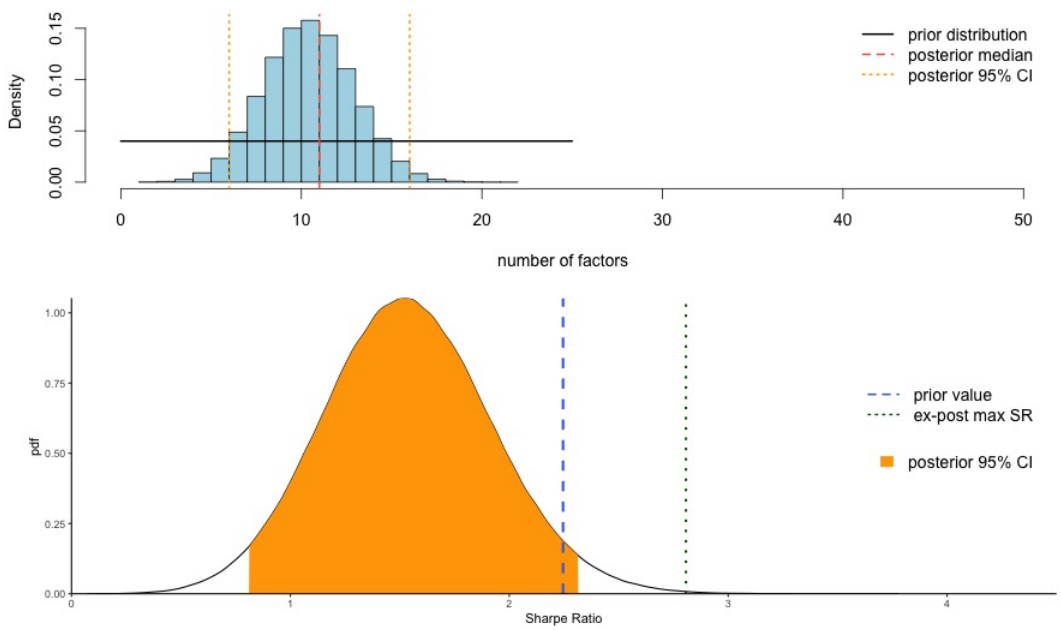


Figure A.4: Posterior SDF dimension and Sharpe Ratio with duration-adjusted returns – bond factor zoo.

Posterior distributions of the number of factors to be included in the SDF (top panel) and of the SDF-implied Sharpe ratio (bottom panel), computed using 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors, with returns computed in excess of the duration-matched U.S. Treasury Bond rate of return ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

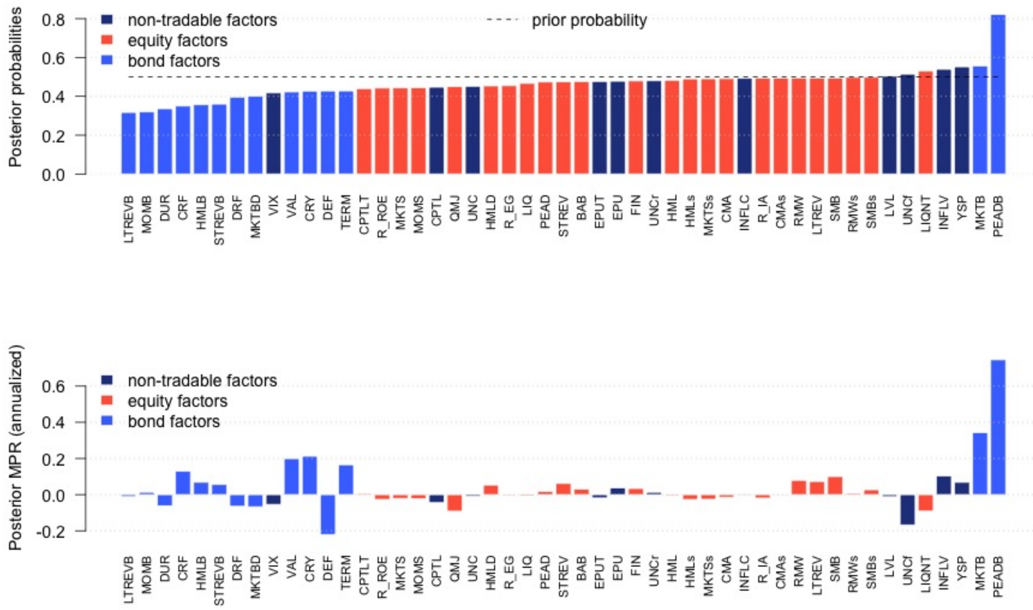


Figure A.5: Posterior factor probabilities and market prices of risk – bond and equity factor zoo.

The figure plots posterior factor probabilities (top panel), $\mathbb{E}[\gamma_j|\text{data}]$, and the corresponding posterior market prices of risk (bottom panel), $\mathbb{E}[\lambda_j|\text{data}]$, computed using the continuous spike-and-slab approach of Section 2 and 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors, with returns computed in excess of the duration-matched U.S. Treasury Bond rate of return ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

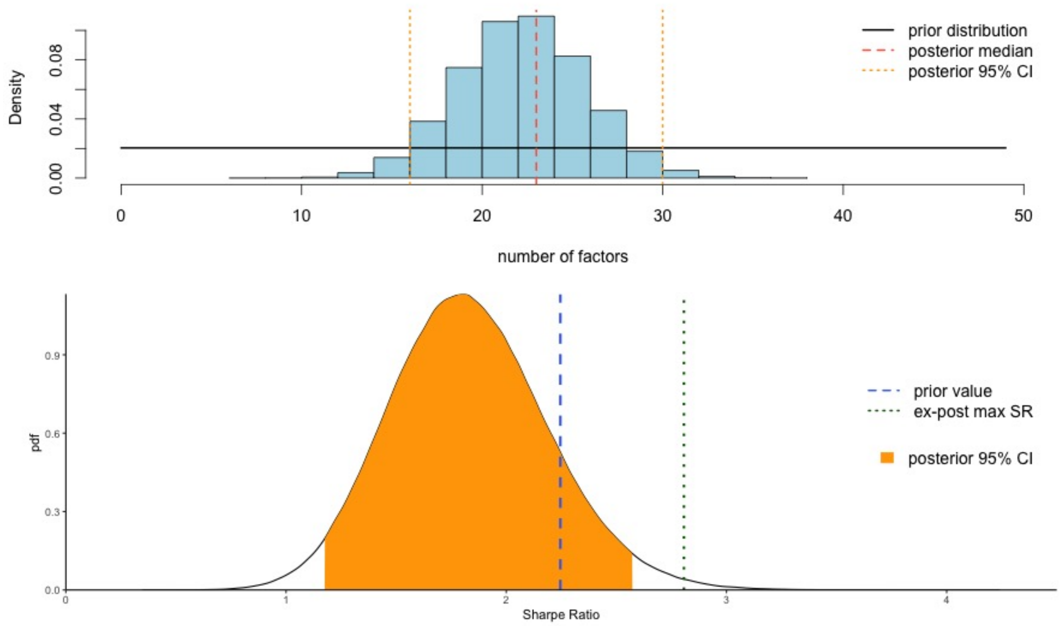


Figure A.6: Posterior SDF dimension and Sharpe Ratio with duration-adjusted returns – bond and equity factor zoo.

Posterior distributions of the number of factors to be included in the SDF (top panel) and of the SDF-implied Sharpe ratio (bottom panel), computed using 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors, with returns computed in excess of the duration-matched U.S. Treasury Bond rate of return ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

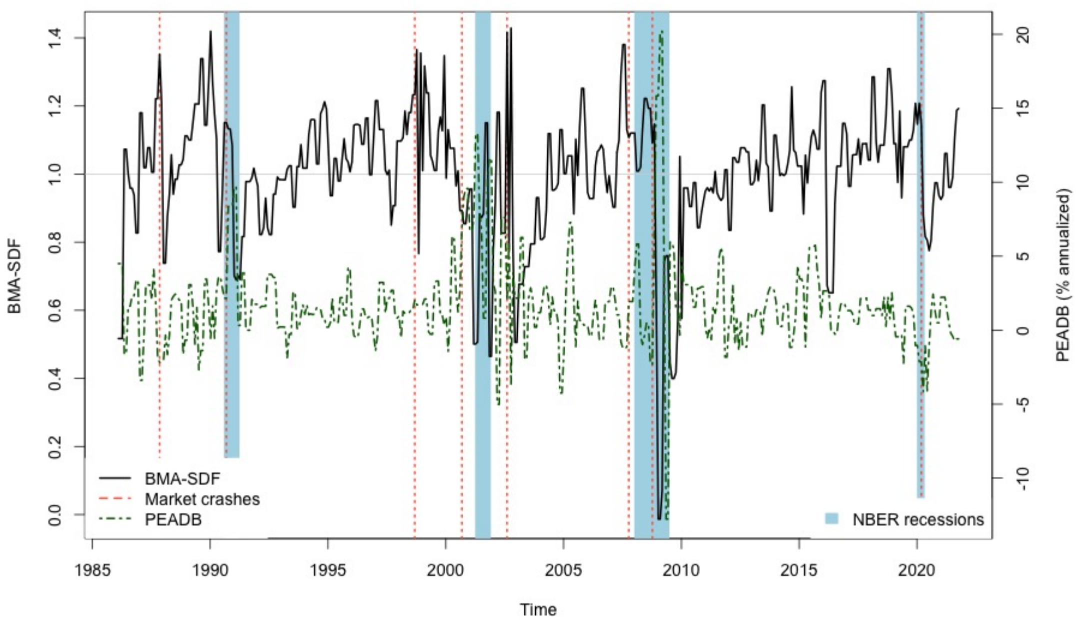
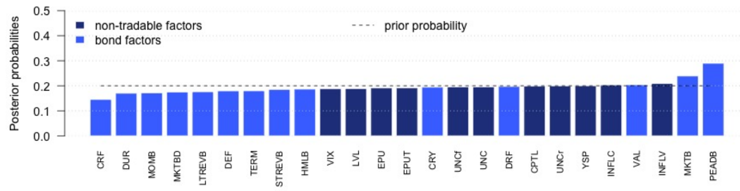


Figure A.7: BMA-SDF, economic cycles, and PEADB with duration-adjusted returns – bond and equity factor zoo.

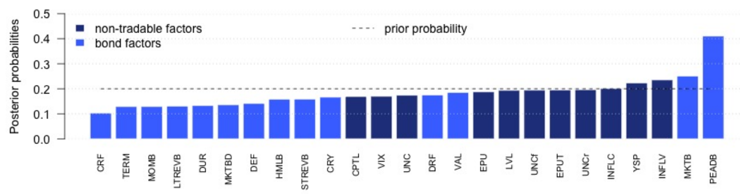
The figure plots the smoothed time series of the posterior mean of the BMA-SDF (left scale), computed using the continuous spike-and-slab approach of Section 2 and 49 bond and equity factors described in Appendix B, and of the smoothed PEAD factor annualized returns (right scale). The blue shaded areas represent NBER-dated recessions, and the red dotted vertical lines correspond to the major stock market crashes identified in Mishkin and White (2002) plus the 2008 and 2020 contractions. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, and the 14 traded bond factors, with returns computed in excess of the duration-matched U.S. Treasury Bond rate of return ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 80% of the expost maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

D Additional figures

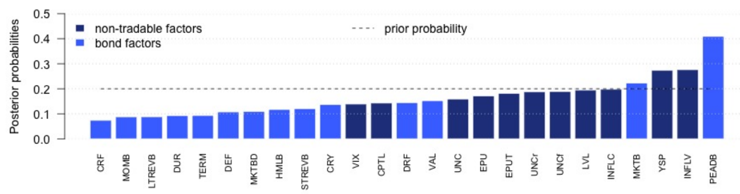
Figure A.8: Posterior factor probabilities in sparse models for different prior SR – bond factor zoo.



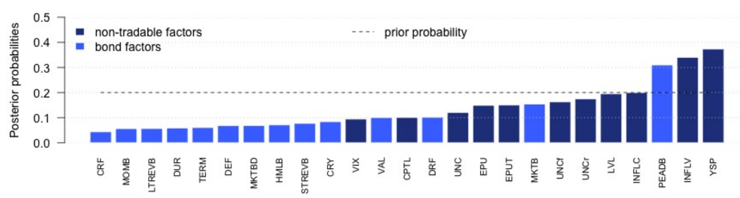
(a) prior Sharpe ratio = 20%



(b) prior Sharpe ratio = 40%



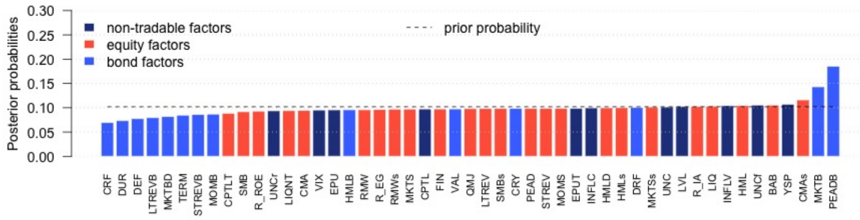
(c) prior Sharpe ratio = 60%



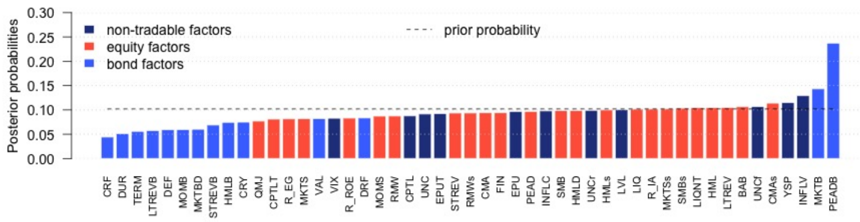
(d) prior Sharpe ratio = 80%

Posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ of 25 bond factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). Prior distribution for the j^{th} factor inclusion is a Beta(3, 12), yielding a 0.2 prior mean for γ_j. Prior Sharpe ratio set to 20%, 40%, 60% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

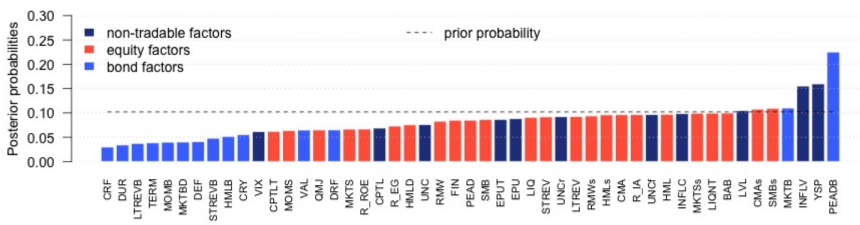
Figure A.9: Posterior factor probabilities in sparse models for different prior SR – bond and equity factor zoo.



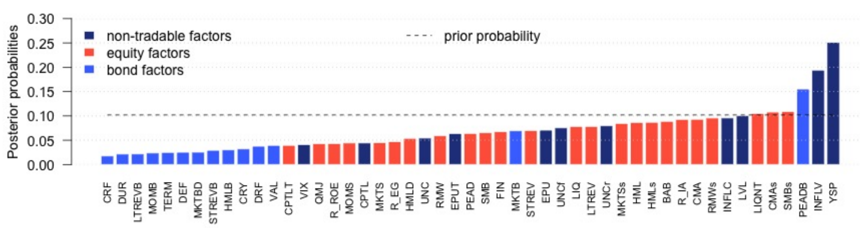
(a) prior Sharpe ratio = 20%



(b) prior Sharpe ratio = 40%



(c) prior Sharpe ratio = 60%



(d) prior Sharpe ratio = 80%

The figure plots posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ computed using the continuous spike-and-slab approach of Section 2 and 25 bond factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(3.45, 30.7), yielding a ≈ 0.10 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, 60%, and 80% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

Internet Appendix for:

The Corporate Bond Factor Zoo

Abstract

This Internet Appendix provides additional tables, figures, and results supporting the main text.

IA.1 Detailed data and variables construction

The following sections describe the various databases that we use in the paper. Across all databases, we filter out bonds which have a time-to-maturity of less than 1-year. Furthermore, for consistency, across all databases, we define bond ratings as those provided by Standard & Poors (S&P). We include the full spectrum of ratings (AAA to D), but exclude bonds which are unrated. For each database that we consider, we (the authors) *do not* winsorize or trim bond returns in any way.

IA.1.1 Corporate bond databases

IA.1.1.1 Mergent Fixed Income Securities Database (FISD) database

Mergent Fixed Income Securities Database (FISD) for academia is a comprehensive database of publicly offered U.S. bonds, research market trends, deal structures, issuer capital structures, and other areas of fixed income debt research. We apply to the FISD data the standard filters used in the previous literature:

1. Only keep bonds that are issued by firms domiciled in the United States of America, `COUNTRY_DOMICILE == 'USA'`.
2. Remove bonds that are private placements, `PRIVATE_PLACEMENT == 'N'`.
3. Only keep bonds that are traded in U.S. Dollars, `FOREIGN_CURRENCY == 'N'`.
4. Bonds that trade under the 144A Rule are discarded, `RULE_144A == 'N'`.
5. Remove all asset-backed bonds, `ASSET_BACKED == 'N'`.
6. Remove convertible bonds, `CONVERTIBLE == 'N'`.
7. Only keep bonds with a fixed or zero coupon payment structure, i.e., remove bonds with a floating (variable) coupon, `COUPON_TYPE != 'V'`.
8. Remove bonds that are equity linked, agency-backed, U.S. Government, and mortgage-backed, based on their `BOND_TYPE`.

9. Remove bonds that have a "non-standard" interest payment structure or bonds not caught by the variable coupon filter (COUPON_TYPE). We remove bonds that have an INTEREST_FREQUENCY equal to −1 (N/A), 13 (Variable Coupon), 14 (Bi-Monthly), and 15 and 16 (undocumented by FISD). Additional information on INTEREST_FREQUENCY is available on page 60 to 67 of the FISD Data Dictionary 2012 document.

IA.1.1.2 Bank of America Merrill Lynch (BAML) database

The BAML data is made available by the Intercontinental Exchange (ICE) and provides daily bond price quotes, accrued interest, and a host of pre-computed corporate bond characteristics such as the bond option-adjusted credit spread (OAS), the asset swap spread, duration, convexity, and bond returns in excess of a portfolio of duration-matched Treasuries. The ICE sample spans the time period 1997:01 to 2021:09 and includes constituent bonds from the ICE Bank of America High Yield (H0A0) and Investment Grade (C0A0) Corporate Bond Indices.

ICE bond filters. We follow van Binsbergen, Nozawa, and Schwert (2023) and take the last quote of each month to form the bond-month panel. We then merge the ICE data to the filtered Mergent FISD database. The following ICE-specific filters are then applied:

1. Only include corporate bonds, `Ind_Lvl_1 == 'corporate'`
2. Only include bonds issued by U.S. firms, `Country == 'US'`
3. Only include corporate bonds denominated in U.S. Dollars, `Currency == 'USD'`

BAML/ICE bond returns. Total bond returns are computed in a standard manner in ICE, and no assumptions about the timing of the last trading day of the month are made because the data is quote based, i.e., there is always a valid quote at month-end to compute a bond return. This means that each bond return is computed using a price quote at exactly the end of the month, each and every month. This introduces homogeneity into the bond returns because prices are sampled at exactly the same time each month. ICE only provides bid-side pricing, meaning bid-ask bias is inherently not present in the monthly sampled prices, returns and credit spreads. The monthly ICE return variable is (as denoted in the original database) `trr_mtd_loc`, which is the month-to-date return on the last business day of month t.

IA.1.1.3 Lehman Brothers (LHM) database

The Lehman Brothers Bond database holds monthly price data for corporate (and other) bonds from January 1973 to December 1997. The database categorizes the prices as either quote or matrix prices and identifies whether the bonds are callable or not. However, as per the findings of Chordia, Goyal, Nozawa, Subrahmanyam, and Tong (2017), the difference between quote and matrix prices or callable and non-callable bonds does not have a material impact on cross-sectional return predictability. Hence, we include both types of observations. In addition, the Lehman Brothers data provides key bond details such as the amount outstanding, credit rating, offering date, and maturity date. For the main results, we use the LHM data from 1986:01 to 1996:12.

LHM filters. As for the other databases, we merge the LHM data to the pre-filtered Mergent database and then apply the following LHM-specific filters following Elkamhi, Jo, and Nozawa (2023):

1. Only include corporate bonds classified as ‘industrial’, ‘telephone utility’, ‘electric utility’, ‘utility (other)’, and ‘finance’, as per the LHM industry classification system, `icode == {3 | 4 | 5 | 6 | 7}`.
2. Remove the following dates for which there are no observations or valid return data, `date == {1975-08 | 1975-09 | 1984-12 | 1985-01}`.

LHM returns. The LHM bond database includes corporate bond returns that have been pre-computed. The accuracy of the LHM return computation has been verified empirically by Elkamhi, Jo, and Nozawa (2023).

LHM additional filters. We follow Bessembinder, Kahle, Maxwell, and Xu (2008) and Chordia, Goyal, Nozawa, Subrahmanyam, and Tong (2017) and apply the following filters to the LHM data to account for potential data errors:

1. Remove observations with large return reversals, defined as a 20% or greater return followed by a 20% or greater return of the opposite sign.

2. Remove observations if the prices appear to bounce back in an extreme fashion relative to preceding days. Denote R_t as the month t return, we exclude an observation at month t if $R_t \times R_{t-k} < -0.02$ for $k = 1, \ldots, 12$.

3. Remove observations if prices do not change for more than three months, i.e., $\frac{P_t}{P_{t-3}} - 1$ `!= 0`, where P is the quoted or matrix price.

IA.1.2 Combined data

For our main results, we rely on the data set that combines the LHM, and ICE data sets over the sample period 1986:01–2021:09. The data is spliced together as follows:

1. From 1986:01–1996:12 we use the LHM data.

2. From 1997:01–2021:09 we use the ICE data.

IA.1.3 Robustness – WRDS bond database

The Wharton Research Data Services (WRDS) Bond Database is a pre-processed monthly bond data set that uses the Enhanced Trade Reporting and Compliance Engine (TRACE) and Mergent FISD bond databases. It was introduced by WRDS in April 2017. The data is publicly available (requires a valid subscription to WRDS). After logging in to WRDS, the data is available here. We use the version of the WRDS data set that spans the sample 2002:09–2021:09.

WRDS bond returns. The WRDS data team provides us with three different bond return variables: `RET_EOM` (returns are computed using bond prices that land on any day of the month), RET_L5M (a bond must trade on the last five days of the month), and RET_LDM (a bond must trade on the last day of the month). For the results based on the WRDS Bond Database, we always use RET_L5M, i.e., a return is valid if the bond trades on the last five days of month t and month $t-1$. However, the publicly available data set any bond return that is greater than 100% to 100%, i.e., returns are truncated at this level. Although this does not make any material difference whatsoever to the main results, we carefully address the issue below.

WRDS bond returns truncation correction. We carefully adjust for the truncation of bonds with returns greater than +100% imposed by WRDS, by setting any bond return which is truncated to the return observed in the ICE database, i.e., if the WRDS bond return is equal to 100% (truncated), we set this value to the bond return from ICE as the 'true' bond return. If the ICE return is missing, we set the value to the return computed from the TRACE data itself. These adjustments do not make any material difference to the robustness results. In total we identify only 94 cases where the truncation occurs, and we are able to address 91 of them. The remaining 3 cases are removed.

WRDS bond filters. To align the data to the Bank of America Merrill Lynch (BAML) corporate bond database provided by the Intercontinental Exchange (ICE), we follow Andreani, Palhares, and Richardson (2023) and use the following filters (all using data provided by WRDS):

1. Remove investment (IG) rated bonds that have less than USD 150 million outstanding prior to, and including, November 2004, and less than USD 250 million after November 2004.

2. Remove non-investment grade (HY) rated bonds that have less than USD 100 million outstanding prior to, and including, September 2016, and less than USD 250 million after September 2016.

3. Remove bonds which are classified as zero-coupon, `bond_type == 'CMTZ'`.

4. Remove bonds which are classified as convertible, `conv == 'N'`.

We merge the WRDS data to the Mergent FISD database (also publicly available via the WRDS data platform) and apply the filters already discussed above. This procedure delivers a transaction-based TRACE data set that closely aligns to the quote-based ICE data.

IA.1.4 Correcting price-based TRACE characteristics for microstructure noise

As first emphasized by Bartram, Grinblatt, and Nozawa (2020), BGN, price measurement error shared by a month-end transaction 'price-based' signal and the subsequent return generates correlation between the two. This affects, for example signals based on bond credit spreads, yields and size (market capitalization) for the TRACE (WRDS) database. We follow the methodology of BGN and Dickerson, Robotti, and Rossetti (2023), DRR, and define the 'month-end' price-based signal to use a transaction-price at least one-business day before the price used to compute a month-end ex ante return. This methodology dampens the transmission mechanism of market microstructure noise (MMN) inherent in the price-based TRACE price signals. DRR show that by accounting for the transmission of the measurement error in this manner, the out-of-sample TRACE price-based anomalies are aligned to those observed when using the ICE data set.

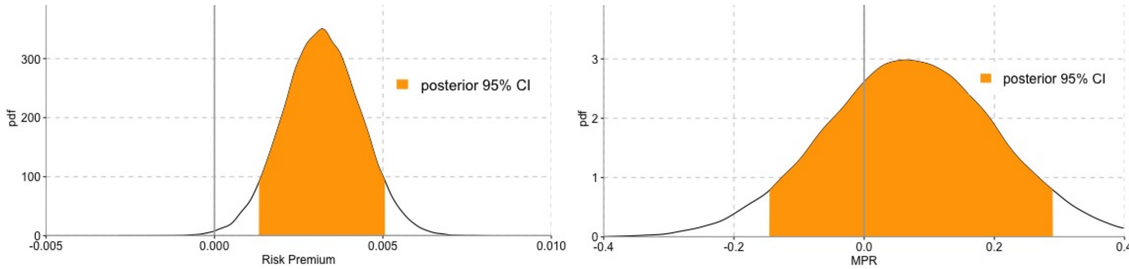
IA.2 CAPMB: Two-pass regression risk premium vs. SDF-based market price of risk

In this section we report two-pass regression estimates of the risk premium attached to MKTB as sole factor as well as linear SDF estimates of the market price of risk in the CAPMB model.

To understand why the two types of estimations can lead to very different outcomes, let's consider a simple example with two (demeaned) tradable risk factors only, i.e. $\boldsymbol{f}_t = [f_{1,t}, f_{2,t}]^\top$, and suppose for simplicity that their covariance matrix is

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Figure IA.1: CAPMB: two pass-regression risk premium, and market price of risk



(a) Two-pass regression risk premium (b) SDF-based market price of risk

The figure plots the posterior distributions of the two-pass regression ex post risk premium, left panel, and SDF-based market price of risk, right panel, of a model with MKTB as the only risk factor, i.e. CAPMB. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate. The prior Sharpe ratio does not impose any shrinkage being set to the ex post Sharpe ratio of the MKTB factor. Sample period: 1986:01 to 2021:09 ($T = 429$).

Suppose further that only the first factor is part of the SDF, and has a market price of risk equal to κ. That is

$$M_t = 1 - \boldsymbol{f}_t^\top \boldsymbol{\lambda_f} = 1 - [f_{1,t}, f_{2,t}]^\top \begin{bmatrix} \kappa \\ 0 \end{bmatrix} = 1 - f_{1,t}\kappa$$

Denoting with $\boldsymbol{\mu}_{RP} = \left[\mu_{RP,1}, \mu_{RP,2}\right]^\top$ the vector of risk premia of the factors, applying the fundamental asset pricing equation to the returns generated by the factors we have

$$\boldsymbol{\mu}_{RP} = \Sigma \boldsymbol{\lambda_f} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} \kappa \\ 0 \end{bmatrix} = \begin{bmatrix} \kappa \\ \rho\kappa \end{bmatrix}.$$

That is, the second factor, that is *not* part of the SDF, commands nevertheless a non-zero risk premium (equal to $\rho\kappa$) as long as the factor has non-zero correlation (i.e., as long as $\rho \neq 0$) with the true risk factor—the one that is part of the SDF . This also implies that a two-pass regression method that uses the second factor as the sole driver of a cross-section of asset returns will estimate its ex post risk premium as being non-zero – in fact, the estimated risk premium for the second factor will be inflated relative to its true value. This is due to the fact that the estimated betas of f_2 will be, in population, smaller than the ones of f_1 by a factor equal to ρ. Hence, in population, the two pass regression will yield an estimated risk premium for f_2 equal to $\rho^{-1}\kappa$ (where $|\rho| \leq 1$).

To estimate the SDF of the CAPMB model we rely on the Bayesian-SDF estimator in Definition 1 of Bryzgalova, Huang, and Julliard (2023). This is equivalent to the method presented in Section 2 under the null that MKTB is the only factor in the SDF with probability 1 and that the model is true. To put the comparison of MRP and ex post risk premia estimates on the same footing, we estimate the two pass regression using the Bayesian implementation of the Fama and MacBeth (1973) method in Bryzgalova, Huang, and Julliard (2022).[*] Posterior distributions of the two-pass regression ex post risk premium and SDF-based market price of risk are plotted, respectively, in panels (a) and (b) of Figure IA.1. The estimates suggests that, albeit MKTB carries a sizable and significant risk premium, it is very unlikely that the data are generated by a "true" latent SDF with MKTB as the only factor—the (Bayesian) p-value of its market price of risk being equal to zero is about 64.78%.

[*]In particular, to implement the two-pass regression we use the fact that the posterior of the time series estimates of the MKTB betas and average excess returns of the test assets, $\boldsymbol{\mu}_R$, under a flat prior, follow the canonical Normal-inverse-Wishart posterior of a linear regression model (see, e.g., Bauwens, Lubrano, and Richard (1999)), and can be sampled accordingly. Furthermore, conditional on the posterior draws of the first-pass regression, and under the null of the model being true, the posterior distribution of the (OLS) second-pass regression is a Dirac distribution at $(\beta^\top\beta)^{-1}\beta^\top\boldsymbol{\mu}_R$, where $\beta = (\mathbf{1}_N, \beta_f)$, $\beta_f = \boldsymbol{C}_f\Sigma_f^{-1}$ and the factor f is simply MKTB.

IA.3 Additional figures and tables

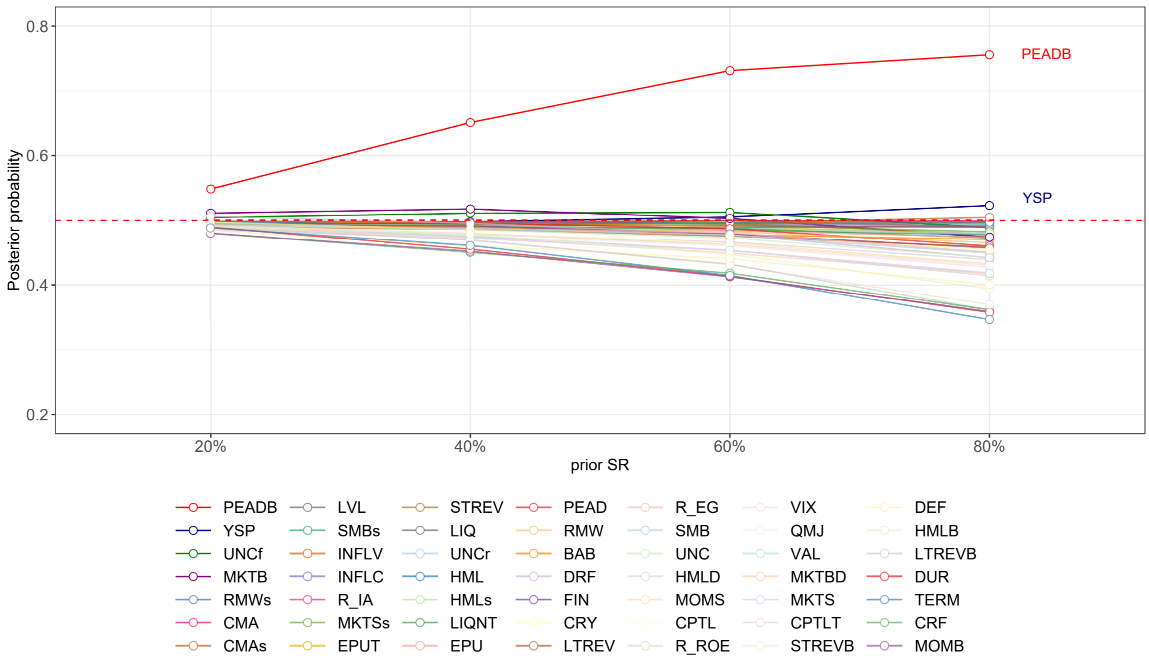


Figure IA.2: Posterior factor probabilities – bond and stock factor zoo – WRDS TRACE sample.

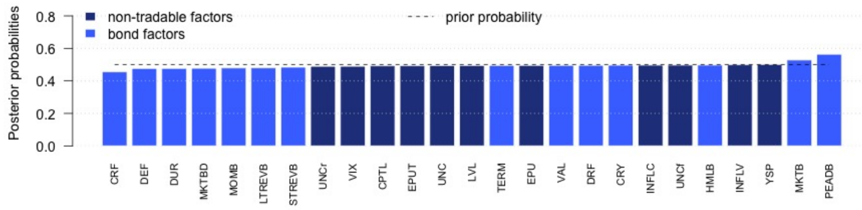
Posterior probabilities, $\mathbb{E}[\gamma_j|\text{data}]$, of 49 bond and stock factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). Both the test assets and the 14 traded bond factors are computed with the WRDS TRACE data. The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. Posterior probabilities for different values of the prior Sharpe ratio, $\sqrt{\mathbb{E}_\pi[SR_{\boldsymbol{f}}^2 \mid \sigma^2]}$, set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. Sample period: 2002:09 to 2021:09 ($T = 229$).

Table IA.I: Cross-sectional asset pricing – WRDS TRACE sample.

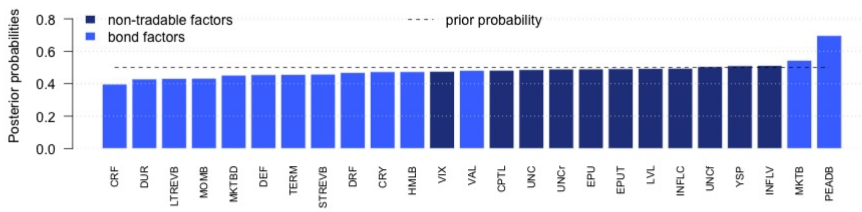
	In-sample				Out-of-sample			
	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}	RMSE	MAPE	R^2_{OLS}	R^2_{GLS}
Panel A: BMA-SDF with 25 bond factors (33.6 mn models)								
prior SR = 20%	0.180	0.129	0.282	0.064	0.121	0.093	0.104	0.037
prior SR = 40%	0.146	0.109	0.523	0.150	0.116	0.090	0.185	0.068
prior SR = 60%	0.125	0.098	0.652	0.217	0.115	0.089	0.195	0.081
prior SR = 80%	0.114	0.090	0.712	0.259	0.115	0.089	0.201	0.092
Panel B: BMA-SDF with 49 bond and stock factors (563 tn models)								
prior SR = 20%	0.187	0.133	0.220	0.049	0.123	0.094	0.074	0.029
prior SR = 40%	0.158	0.116	0.448	0.125	0.117	0.091	0.160	0.058
prior SR = 60%	0.134	0.103	0.598	0.202	0.115	0.090	0.191	0.076
prior SR = 80%	0.118	0.093	0.693	0.278	0.114	0.089	0.210	0.089
Panel C: Benchmark models and most likely factors								
CAPMB	0.195	0.113	0.152	0.071	0.110	0.083	0.264	0.067
CAPM	0.305	0.245	−1.069	−0.031	0.165	0.120	−0.663	0.011
FF5	0.201	0.147	0.101	0.043	0.135	0.102	−0.113	0.025
PEADB	0.318	0.220	−1.242	0.087	0.150	0.117	−0.362	−0.068
Top factors bond	0.126	0.103	0.649	0.366	0.133	0.101	−0.081	0.073
Top factors all	0.154	0.133	0.470	0.417	0.134	0.106	−0.101	0.034

In-sample and cross-sectional out-of-sample pricing performance of BMA-SDF, notable factor models, and factors with a posterior probability greater than 50%. We use the WRDS TRACE data spanning the sample period 2002:09 to 2021:09 ($T = 229$). We use GMM-GLS to estimate factor risk prices for CAPMB, CAPM, and the and Fama and French (1992, 1993) model, which includes the MKTS, SMB, HML, DEF and TERM factors, and a single-factor model with PEADB. The 'Top factors bond' model includes PEADB, YSP, UNCf, MKTB and DRF. The 'Top factors all' model includes additionally CMAs. For the BMA-SDF, we report results for a range of prior Sharpe ratio values that are set as 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio of the 64 bond portfolios and traded factors. In-sample (IS) test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity and the 14 traded bond factors ($N = 64$). Out-of-sample (OS) test assets assets include decile-sorted portfolios on bond historical value-at-risk (95%), duration, bond value, bond book-to-market, long-term reversals, momentum and the 17 Fama French industry portfolios ($N = 77$). In cross-sectional OS tests, models are first estimated using the baseline IS test assets and then used to price (with no additional parameter estimation) the OS assets. All data is standardized, that is, pricing errors are in Sharpe ratio units.

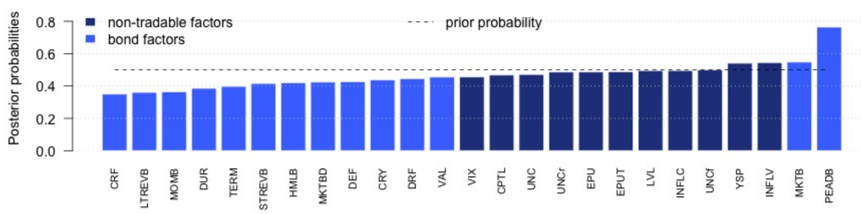
Figure IA.3: Posterior factor probabilities for different prior SR – bond factor zoo.



(a) prior Sharpe ratio = 20%



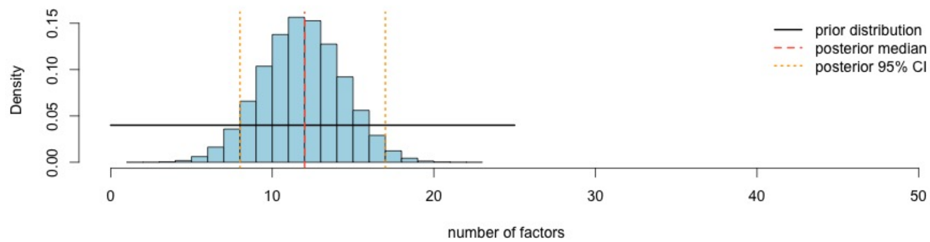
(b) prior Sharpe ratio = 40%



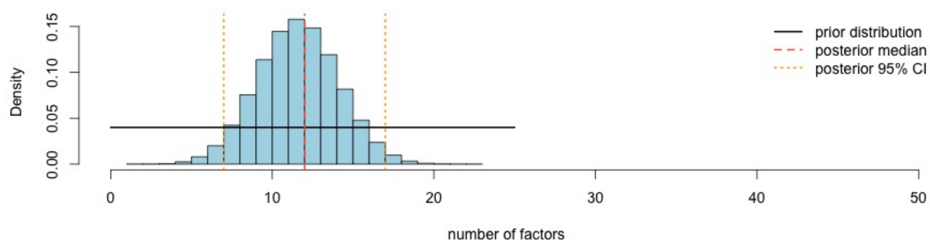
(c) prior Sharpe ratio = 60%

The figure plots posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ computed using the continuous spike-and-slab approach of Section 2 and 25 bond factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

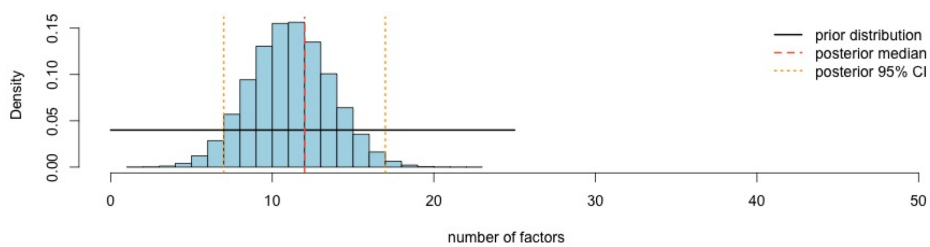
Figure IA.4: Posterior SDF dimensionality for different prior SR – bond factor zoo.



(a) prior Sharpe ratio = 20%



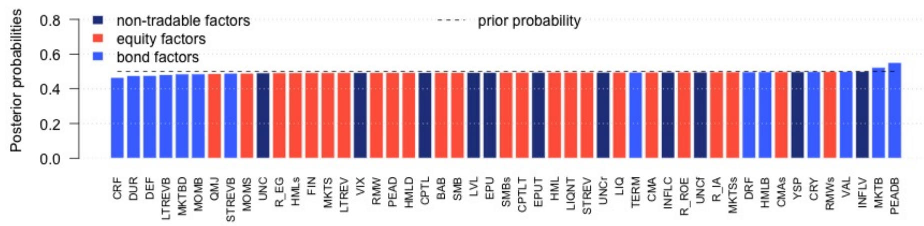
(b) prior Sharpe ratio = 40%



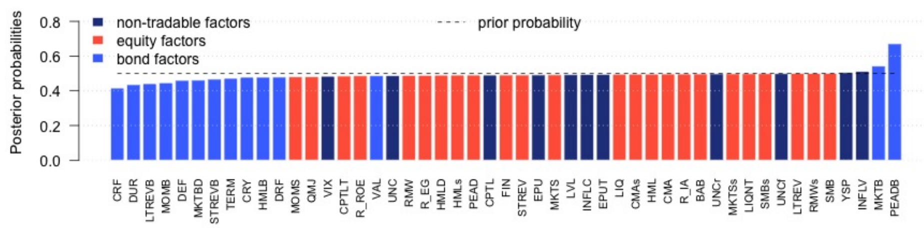
(c) prior Sharpe ratio = 60%

Posterior distributions of the number of factors to be included in the SDF, for different values of the prior Sharpe ratio, computed using 25 bond related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

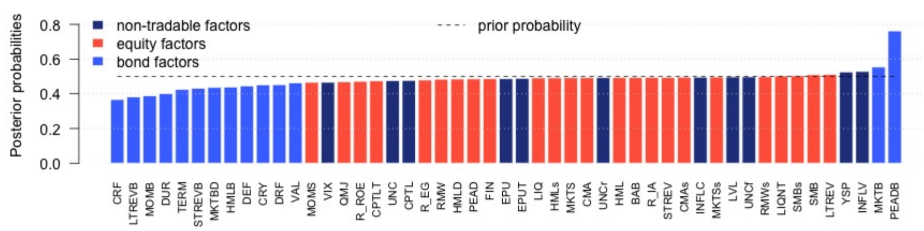
Figure IA.5: Posterior factor probabilities for different prior SR – bond and equity factor zoo.



(a) prior Sharpe ratio = 20%



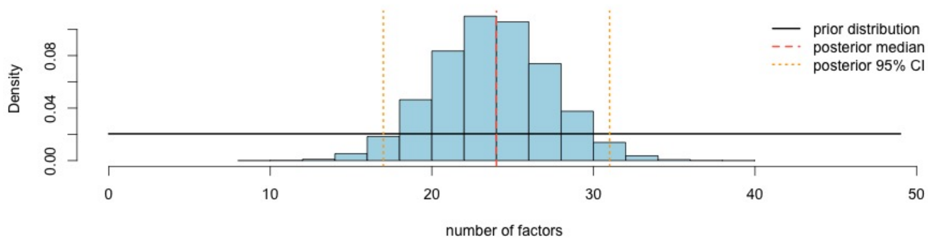
(b) prior Sharpe ratio = 40%



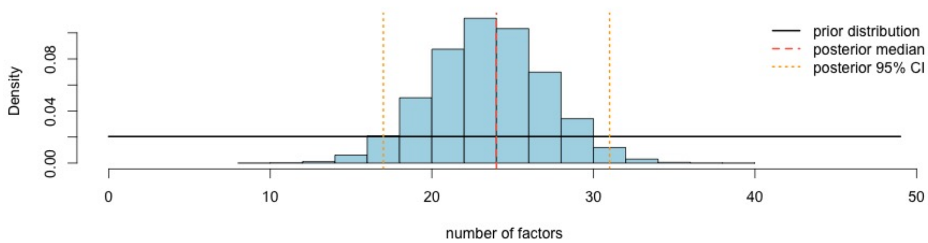
(c) prior Sharpe ratio = 60%

The figure plots posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ computed using the continuous spike-and-slab approach of Section 2 and 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

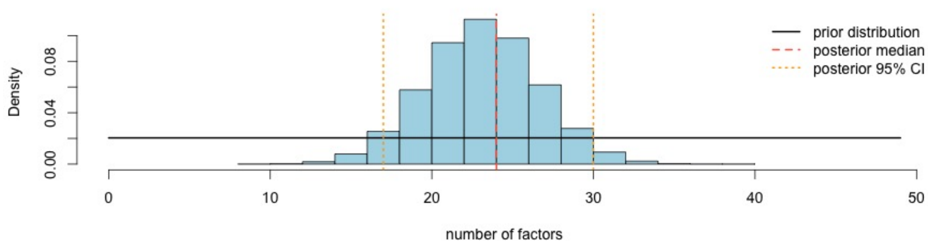
Figure IA.6: Posterior SDF dimensionality for different prior SR – bond and equity factor zoo.



(a) prior Sharpe ratio = 20%



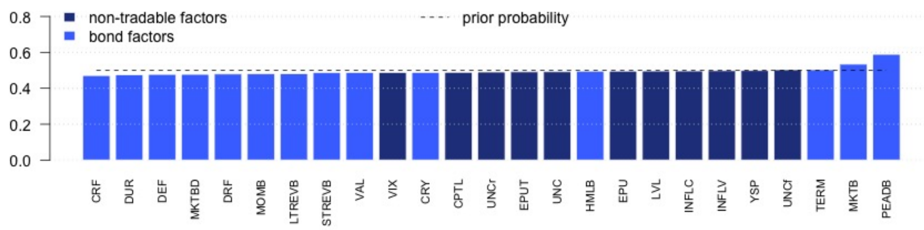
(b) prior Sharpe ratio = 40%



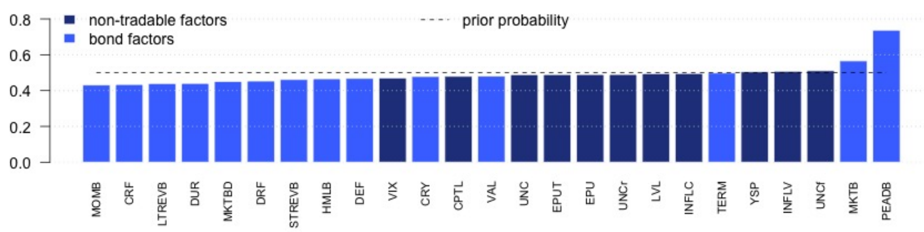
(c) prior Sharpe ratio = 60%

Posterior distributions of the number of factors to be included in the SDF, for different values of the prior Sharpe ratio, computed using 49 bond and equity related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

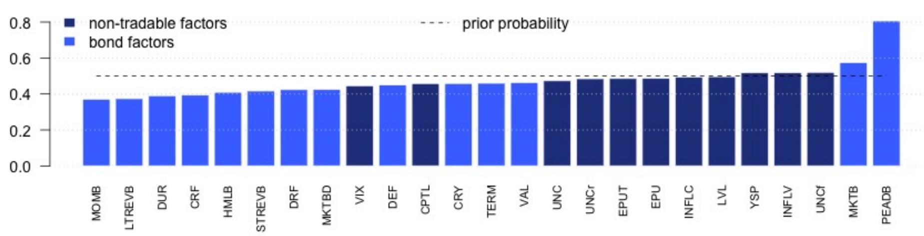
Figure IA.7: Posterior factor probabilities for different prior SR with duration adjusted returns – bond factor zoo.



(a) prior Sharpe ratio = 20%



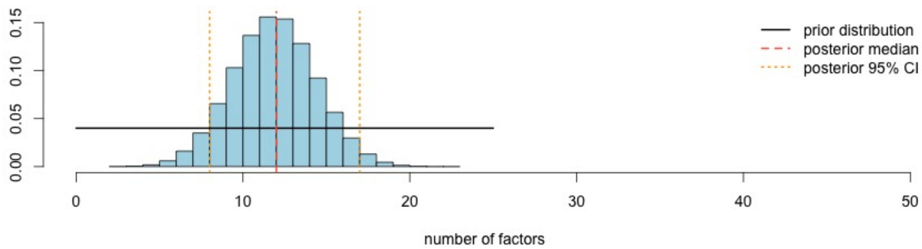
(b) prior Sharpe ratio = 40%



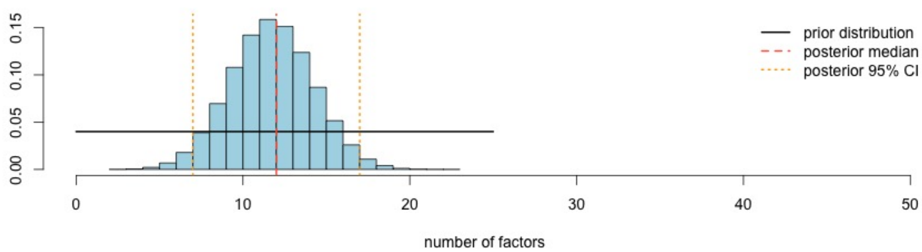
(c) prior Sharpe ratio = 60%

The figure plots posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ computed using the continuous spike-and-slab approach of Section 2 and 25 bond factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

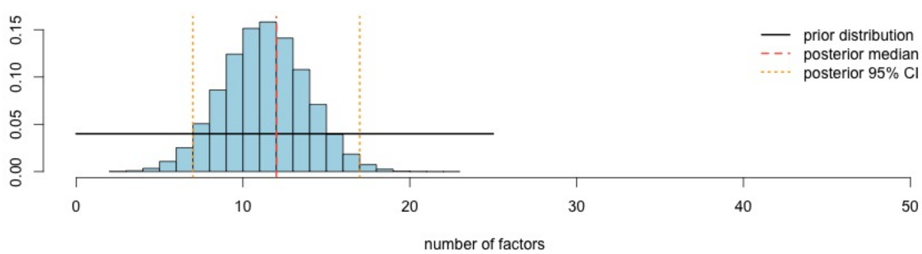
Figure IA.8: Posterior SDF dimensionality for different prior SR with duration-adjusted returns – bond and equity factor zoo.



(a) prior Sharpe ratio = 20%



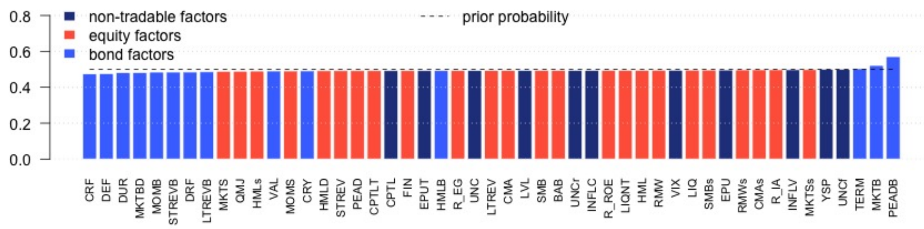
(b) prior Sharpe ratio = 40%



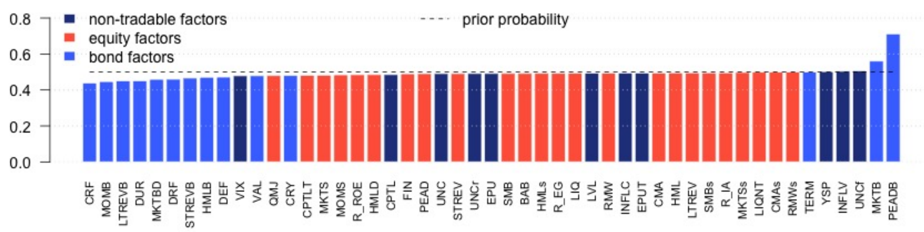
(c) prior Sharpe ratio = 60%

Posterior distributions of the number of factors to be included in the SDF, for different values of the prior Sharpe ratio, computed using 49 bond and equity related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, 60% and 80% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

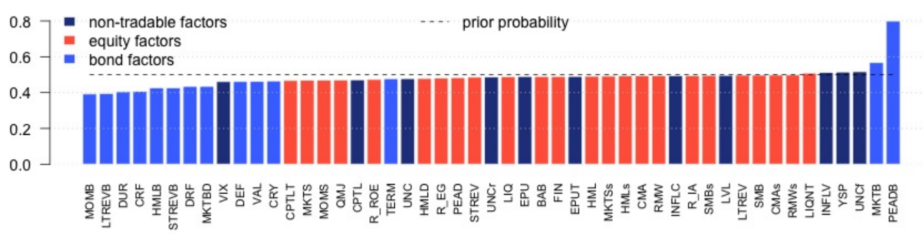
Figure IA.9: Posterior factor probabilities for different prior SR with duration-adjusted returns– bond and equity factor zoo.



(a) prior Sharpe ratio = 20%



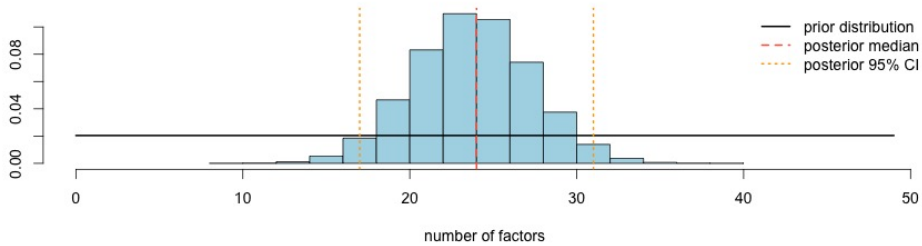
(b) prior Sharpe ratio = 40%



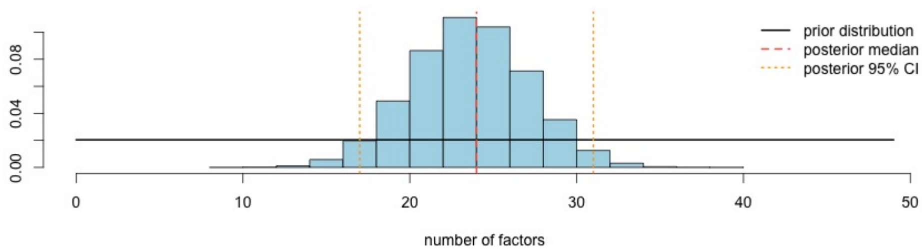
(c) prior Sharpe ratio = 60%

The figure plots posterior factor probabilities, $\mathbb{E}[\gamma_j|\text{data}]$ computed using the continuous spike-and-slab approach of Section 2 and 49 bond and equity factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).

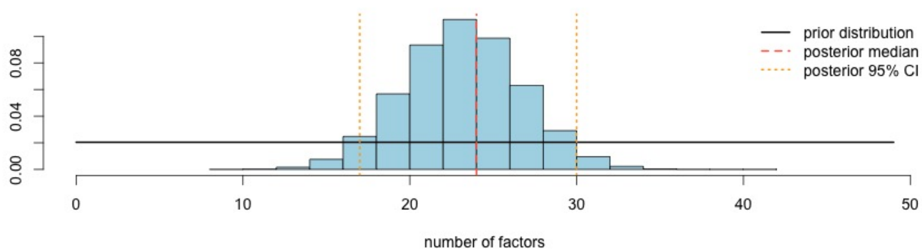
Figure IA.10: Posterior SDF dimensionality for different prior SR with duration-adjusted returns – bond and equity factor zoo.



(a) prior Sharpe ratio = 20%



(b) prior Sharpe ratio = 40%



(c) prior Sharpe ratio = 60%

Posterior distributions of the number of factors to be included in the SDF, for different values of the prior Sharpe ratio, computed using 49 bond and equity related factors described in Appendix B. Test assets include 50 bond portfolios sorted on credit spreads, size, rating and maturity, with returns computed in excess of the one-month risk-free rate, and the 14 traded bond factors ($N = 64$). The prior distribution for the j^{th} factor inclusion is a Beta(1, 1), yielding a 0.5 prior expectation for γ_j. The prior Sharpe ratio is set to 20%, 40%, and 60% of the ex post maximum Sharpe ratio achievable with the 64 bond portfolios and traded factors. Sample period: 1986:01 to 2021:09 ($T = 429$).