Asset Pricing, not Stock Pricing

Chunjie Wang*

This Version: March 22, 2024

Abstract

I construct stochastic discount factors (SDFs) from unlevered asset returns and a large number of cross-sectional stock return predictors. I adopt a Bayesian approach in this high-dimensional setting and obtain a characteristics-sparse 5-factor SDF that summarizes 62.4% of out-of-sample cross-sectional variations. Implied mean-variance efficient portfolio outperforms its optimal levered counterparty that consists of 88 factors in both out-of-sample Sharpe ratio and market alphas. Unlevered market returns eliminate the alpha of the efficient stock returns. The failure to shrink the number of factors in recent literature is partly due to constructing factor-mimicking portfolios in the suboptimal levered return space.

*chunjie.wang@phdstudent.hhs.se

Introduction

All factor models in empirical asset pricing are derived as specilizations of the comsumption-based model (Cochrane, 2009). It is often forgotten that most factors are merely intuitively motivated proxies for the marginal rate of substitution, or the stochastic discount factor $M_{t+1} = \beta u'(c_{t+1})/u'(c_t)$ (SDF), rather than risks themselves. Specifically, any characteristics that predict asset returns or aggregate macroeconomic risks can be defended as pricing factors because they affect consumptions and thereby marginal utility, granting "fishing license"(Fama, 1991) that leads to "the factor zoo"(Feng, Giglio, and Xiu, 2020). However, two important assumptions bridging the consumption- and characteristics-based models are often made without scrutiny: 1. economically, factor-mimicking portfolios are built in the return space that faithfully reflects the economic theory behind the characteristics as return predictors; and 2. statistically, factors and returns are either bivariate normal.

The violations of the assumptions are prevalent across all asset classes, especially in stock pricing. For instance, the economic foundation for chatacteristics such as size and profitability being stock returns predictors essentially reveal risk profile on a firm level. Hence, the SDF should be projected onto the unlevered return space to fully reflect the economic driver behind. In reality, however, corresponding factors are typically built with stock returns, despite them simply being a derivative to the future cash flow of the firm. What's worse, any defense on the normality and linearity lacks conviction since the return on equities as call options to the underlying assets is a highly non-linear function of the underlying return even under the mildest assumption.

This paper investigates whether the abuse of the assumptions has in part contributed to spurious anomalies in the cross-sectional stock pricing in a high-dimenstional setting.

I adopt a data-driven approach to shrink daily factors built in both levered stock return space and unlevered asset return space (hereafter "levered SDF" and "unlevered SDF") from 100 firm characteristics that existing literature has claimed to have predictive power, compiled by Green, Hand, and Zhang (2017). I unlever stock returns following the structural credit risk model of Merton (1974). To overcome the high-dimensional challenge, I employ an economically motivated Bayesian prior following Kozak, Nagel, and Santosh (2020). The prior distribution relates the first and second moments of candidate factors and disproportionately penalizes coefficients of factors associated with high-eigenvalue pricipal components (PCs). Optimal SDF coefficient estimators are derived under hyperparameters in the posterior that yield the highest out-of-sample (OOS) R^2. This paper uncovers evidence that forcing the assumptions to linear factor models have been overloading the SDF with spurious factors: the optimal unlevered SDF only consists of 5 factors that can explain 62.4% of OOS cross-sectional variation: market, 12-month momentum, return on equity, asset growth, and revenue surpirses. In contrast, the optimal levered SDF consists of 88 factors but only explains 36.9% OOS variations. In addition, I compare the alphas of the mean-variance-efficient (MVE) portfolios implied by the levered and unlevered SDF against various benchmark portfolios. While the annualized alphas of the levered MVE portfolio are significant 5.50% and 3.53% when testing against stock market and Fama-French 4 factors, it drops to an insignificant −1.92% when testing against the unlevered market return. On the other hand, the alphas of the unlevered MVE portfolio are economically higher and statistically more significant than those of the levered MVE portfolio across all benchmarks. To alleviate the concern that Merton-implied portfolios are not directly tradable, I also construct a set of "tradable unlevered factors" that are readily accessible with a combination of stocks and risk-free treasuries and the results persist. This paper points to a parsimonious CAPM-like factor model in the unlevered space with stronger OOS performance.

Admittedly, these assumptions for factors to proxy marginal utility growth are by no means inexorable, especially when strictly adhering to the rules usually introduces statistical challenge and does not necessarily translate to better economic interpretation due to a lack of granular data and measurement errors. Relaxing the assumptions by building factor-mimicking portfolios in the return space of interest is desirable for its enormous practicality: in a naïve OLS setting, MVE portfolios on a day-by-day basis are revealed simulataneously with the coefficient estimation by simply regressing the mean on the covariance matrix of the factors. Therefore, researchers have deliberately chosen to price stocks with stock factors, bonds with bond factors, and so on because they strike a balance between statistical convenience and factors' roles as proxies for marginal utility growth.

However, the balance has tipped in favor of a more rigorous discipline for factor models recently. In face of hundreds of "anomalies", researchers have been searching for a characteristics-sparse SDF representation that is linear in only a few economically motivated factors with the help of recent development in various unconventional techniques that tackle with high-dimensionality such as elastic nets, partial least squares, and principal component regressions. Unfortunately, it seems new cross-sectional signals keep emerging and the list of factors need constant expansion to capture the new evidence. A lengthy SDF incurs higher cost to falsefully build characteristics-managed portfolios in stock return space when the characteristics capture firm-level risk profile because statistical errors accumulate every time a new ill-defined factor enters the list. When one only focuses on a sparse SDF, the benefits of projecting the SDF onto certain return spaces outweigh the expenses of relaxing the assumptions for factors to proxy marginal rate growth. That was the price asset pricers were willing to pay for a long time. In current high-dimensional climate, the accumulation of the imprecision hinders the search for a sparse factor model and is too grave to ignore.

Revisiting the discussion of what are the fundamental risk factors that capture the marginal rate of substitution is particularly relevant in stock pricing, not only because of "the factor zoo", but also because unlevering stock returns is an intuitive and feasible move to address the common concern of almost all stock-level factors that corresponding firm-level factors are arguably better at reflecting full economic interpretation of the risks. Equities are call options to the underlying assets of a firm with the face value of its debts as the strike price. The cross-sectional variation of the unlevered returns and the risk factors behind are non-linearly magnified by the leverage. Studying a pronounced and non-linear derivative to the total cash flow of firms helps identify the issue of current factor models and potential improvement by taking a step back and reconsider the what defines the total wealth of investors. Unlevered factors are a middle ground between the elusive yet fundamental aggregate consumptions and accessible yet ill-defined levered factors. On the contrary, debts are equivalent of shorting put options of the firm's assets and are by construction less volatile. Coupled with the lack of comprehensive day-by-day data and noises, they are not the best instrument to study whether a return space at a higher level is better at hosting factor-mimicking portfolios. Nevertheless, this paper brings a broader implications in pricing other asset classes such as bonds and currencies.

Recent statistics-driven works sidestepped the problem caused by the large number of potential flawed factors by redirecting the attention to "predicting" from "explaining" the cross-sectional variations. For example, Gu, Kelly, and Xiu (2020) utilized various machine learning techniques such as Deep Neural Networks and Random Forests that are designed to tackle with the challenges from non-linearity and high-dimensionality. Feng, Giglio, and Xiu (2020) adopted advanced statistical tools and raised the bar of marginal importance for new factors to be admitted to an already lenthy list of factors. Kozak, Nagel, and Santosh (2020) obtained a sparse SDF by aggregating information of 49 predictors into a small number of principal components. Although the literature has achieved tremendous success in cross-sectional predictions, it seems we are not moving closer to the ultimate question of what are the fundamental sources of risks if we stop distinguishing the economically-founded factors from the spurious ones. We are guanranteed an empirical success to find an *ex post* MVE portfolio in one sample, but it is unlikely to be efficient in the next sample, *ex ante* or *ex post*. On the other hand, an *ex ante* MVE SDF is unlikely to be efficient in any sample, but it is of the core interests of asset pricing. The only solution is to impose some economically driven disciplines that explain why certain predictors should matter. Therefore, this paper also sends a message that a sparse SDF and the possibility to interpret the SDF is not mutually exclusive. Unlike the aforementioned papers, I kept the characteristics intact, and the improvement of the factor model simply comes from unlevering.

The idea that certain stock pricing anomalies might have unintendedly captured the non-linear transformation of unlevered asset returns is not noval (*e.g.*, Berk, Green, and Naik, 1999; Carlson, Fisher, and Giammarino, 2004). Most works focus on one or two specific anomalies and do not account for the non-linearity explicitly. Rather, they simply include the leverage that is the moneyness of the underlying asset in the regression. Doshi et al. (2019) on the other hand, unlevered the stock returns and found that the size effect is weakened, while the value premium and the volatility puzzle virtually disappear. However, their adoption of portfolio sortings and Fama and MacBeth (1973) (hereafter "FMB") regressions as the empirical strategy only permits tests of limited factors. They unlever returns with a scalar computed from book value of debts, omitting the default risk during their main analyses. Addressing these concerns, this paper, to my best knowledge, is the first to jointly test a wide collection of characteristics-based unlevered factors via an economically motivated Bayesian approach and do so by properly adopting the Merton model in the unlevering process. I find evidence that the omission of the leverage effect on returns has bigger implications beyond a couple of anomalies, and asset pricers need to reconsider the shape of the SDF. As discussed, this paper also revisits the stream of literature on "the factor zoo" of stock pricing that usually provide statistical interpretations (Gu, Kelly, and Xiu, 2020; Feng, Giglio, and Xiu, 2020). In contrast, this paper makes an economically motivated attempt to reduce the dimensionality by targeting the non-linearity introduced by the option-theoretic feature of equity. Despite a small number of principal components of levered factors is sufficient to summarize the cross-section (Kozak, Nagel, and Santosh, 2020), an equally sparse unlevered SDF retains unaltered predictors thus offering ease to potential economic interpretations for surviving factors. Last but not least, this paper also contributes to the expanding body of literature within machine learning applications in empirical finance research (*e.g.*, Gu, Kelly, and Xiu, 2020; Bali et al., 2023). While most papers focus on the predictability of security returns, this paper speculates about the shape of SDF via certain machine learning techniques.

The remainder of the paper is structured as follows: Section I develops the theoretical implication of constructing factor-mimicking portfolios in the levered return space while they reveal risks on firms' whole assets. Section II details the Bayesian approach to be employed in characteristics-based SDF estimation. Section III describes the data and the unlevering process of stock returns. Section IV reports and compares the empirical results for levered and unlevered SDFs. Section V concludes.

I. Factor-Mimicking Portfolios in the Unlevered Space

A firm's equities are call options of the firm's underlying assets and the face value of debts is the strike price. Risk premia on all debts, equities, and other derivatives such as equities' options are linked because all claims must earn the same compensation per unit of risk. The relations among these risk premia are highly non-linear thus demanding prudent deliberation when constructing a linear SDF from security returns. In most cases, these securities are merely instruments to the free cash flow of the same assets. Most hypotheses on fundamental risks describe the exposure on a firm level, and these securities are simply auxiliary. The simplest example is CAPM: it is the covariance with respect to the marginal utility thereby the total wealth of an investor that determines the price of an asset. In theory, a comprehensive measure of the total wealth includes stocks, bonds, human capitals, currencies, real estates and so on. However, it is impossible to make a strict empirical definition, nor can we observe such data that are remotely clean enough to test the hypothesis on a frequent basis. Therefore it is conventional to settle with broad-based stock portfolios such as NYSE and S&P500. One needs to realize that this is a compromise and comes at a cost.

To illustrate why it matters, I borrow from the famous Coval and Shumway (2001) that showed expected call returns non-linearly monotonically increase with the strike price on top of underlying returns[1].

Assuming the existence of an SDF that prices all assets with

$$1 = \mathrm{E}\left[M \cdot R\right] \tag{1}$$

where R is the gross return of any asset, and M is the strictly positive SDF. M is high in bad states of the world and low in good states. Denote v as the random variable of the firm's market value of assets on maturity date. If the firm's face value of debt is B,

[1] Some papers in option pricing such as Christoffersen, Heston, and Jacobs (2013) and Schlag and Sichert (2020) on the other hand suggested a hump-shaped relation. Despite their divergence in the assumptions on the shape of SDF, there is consensus that the variation of the expected call return against the strike price is non-linear. Besides, compared with the wide range of moneyness for options written on equity, the financial leverage of a firm usually entails a single "strike" that is deeply in-the-money, thus the descending phase of the hump-shaped curve would be less relevant for this paper. Given the above reasons, I opt for this simple benchmark to clarify the leverage effect on asset returns.

the expected excess equity return is

$$\mathrm{E}[r_e(B)] = \frac{\mathrm{E}[\max(v-B,0)]}{\mathrm{E}[M \cdot \max(v-B,0)]} - 1\,. \tag{2}$$

The derivative of expected excess equity with respect to the face value of debt can be expressed as[2]

$$\frac{\partial \mathrm{E}[r_e(B)]}{\partial B} = \frac{-\mathrm{Cov}\left[\mathrm{E}(M|v), v-B|v>B\right]}{\left(\mathrm{E}\left[\mathrm{E}(M|v)(v-B)|v>B\right]\right)^2}\,. \tag{3}$$

For firm whose market value of assets negatively correlate with the SDF given that the firm is solvent, the derivative is positive and increases with B. I consider the simplest case just to demonstrate the non-linearity of $\mathrm{E}[r_e(B)]$ against B: the firm is perfectly correlated with the market, CAPM holds in the asset market, and the firm stays solvent at debt maturity. I approximate the utility growth rate by a linear function of the market:

$$\begin{aligned} M &= \beta\frac{u'(c_1)}{u'(c_0)} && (4)\\ &\approx 1 - b\left(r_a - \mathrm{E}\left[r_a\right]\right) && (5)\\ &= 1 - b\frac{v-\mathrm{E}\left[v\right]}{V_0} && (6) \end{aligned}$$

where V_0 is the the firm's present market value of assets; r_a is the excess unlevered return; and b_0 is a positive coefficient. We can solve for $b = \mathrm{E}\left[r_a\right]/\mathrm{Var}\left[r_a\right]$ by pricing r_a itself with the SDF. Inserting Equation (6) back to Equation (3) and imposing the solvent condition gives

$$\frac{\partial \mathrm{E}[r_e(B)]}{\partial B} = \frac{V_0}{\left(V_0 - B\right)^2}\mathrm{E}\left[r_a\right]\,. \tag{7}$$

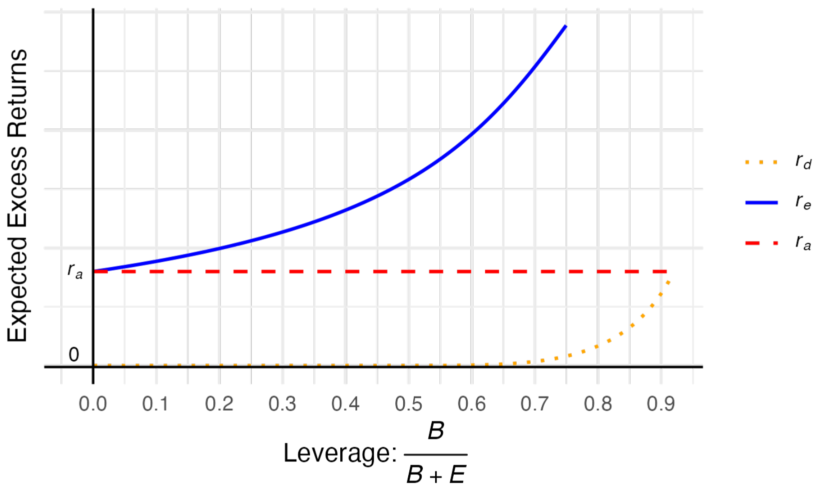
It is evident that Equation (7) is positive and increases in B. Higher debts magnify the required equity return of the firm at an accelerating rate.

[2]Coval and Shumway (2001) or Appendix A provide detailed derivation.

Figure 1 illustrates the relations between expected equity return and financial leverage under this framework. Similar figure is presented in classic textbook such as Berk and DeMarzo (2007) during the discussion of capital structure.

Figure 1: Expected Returns and Leverage

This figure illustrates the general evolvement of the expected excess return of asset (r_a), equity (r_e), and debt (r_d) as a firm levers up. x-axis is the financial leverage, defined as the the ratio of face value of debt (B) over the sum of face value of debt and market capitalization (E). Expected levered equity return r_e monotonically increases in financial leverage and does so at an accelerating rate.



The variation of r_e can be decomposed into systematic unlevered risk and nonlinear representations of leverage risk. The former comes from the covariance between unlevered return r_a (the horizontal line) and the SDF, and manifests on Figure 1 as parallel movements of r_a along the y-axis. However, the latter, shown as the variation of r_e on top of r_a, captures additional compensations for equity holders bearing different leverage risks. If the data generating process follows the unlevered SDF from Equation (5) but a levered SDF $\tilde{M} = 1 - \tilde{b}\,(r_e - \mathrm{E}\,[r_e])$ is assumed, the flawed factor-mimicking portfolio would only capture a fraction of the cross-sectional variation due to the non-linear relationship between r_a and r_e, attributing the rest to spurious anomalies. To see this, since all excess returns in the economy can be priced with the beta representation $\mathrm{E}[r] = \mathrm{Cov}[r, r_a] \cdot b$ with a common price b, it is impossible to find $\tilde{b}$ such that $\mathrm{E}[r] = \mathrm{Cov}[r, r_e] \cdot \tilde{b}$ also hold for all assets due to the non-linearity between r_e and r_a. As a result, an empirical asset pricer might conclude an "anomaly". To add to the challenge, simply including leverage as a covariate does little to capture the leverage effect regardless of the curvature of r_e or whether higher order terms in leverage are included[3], because the relation between leverage and stock returns depends on the sign of the asset risk premium, offsetting the effect in the cross section. Therefore, a correct procedure of return unlevering is crucial in mitigating the omission.

As in most factor models, one might as well make the opposite assumption that the risk factors constituting the SDF are better approximated by the variation of levered return r_e, in which case the unlevered asset would be deemed as the derivative instead. That is, it is not inherently wrong to favor factors in the levered space *per se*. However, neither assumption undermines the argument of non-linear transformation and thereby possible spurious anomalies when the wrong set of returns is chosen to construct a linear factor model. In Section II, I entertain both possibilities by jointly estimating the coefficients and comparing the OOS performance between 100 levered and unlevered characteristics-based factors via the exact same Bayesian approach, in an effort to shed a light on whether returns on asset or equity drive the SDF. In fact, "the true SDF", if exists, is almost certainly a mix in between. That said, I postulate most risks are on the firm-level and both equity and debt holders are exposed indirectly. Many papers discussed possible economic interpretations of stock return anomalies. For example, Carlson, Fisher, and Giammarino (2004) proposed declining growth opportunities as an explanation for neg-

[3]Doshi et al. (2019) in a simulation study found neither leverage or its higher order terms is statistically significant, and the R^2s are very small.

ative relation between size and expected returns. Through this channel, equity holders are only exposed to the secondhand risk of diminishing growth by claiming a piece of the firm's future cash flow.

II. A Bayesian Approach for Factor Discipline

If the expansion of "the factor zoo" is partly due to the leverage effect discussed in Section I, I should find significantly more characteristics-sparse SDF that yields better performance in the unlevered return space than the levered one. This postulation necessitates a regularization technique capable of addressing the joint assessment of numerous factors within a high-dimensional context, while also accommodating filters to allow sparsity. Traditional portforlio sortings or FMB regressions struggle to accomplish these tasks. Therefore, I employ the economically motivated Bayesian approach by Kozak, Nagel, and Santosh (2020) (hereafter "KNS") to regularize the fitting procedure. In this section, I first lay the foundation of asset pricing with characteristics-based factors[4] and introduce the Bayesian approach, then discuss hypotheses under the framework.

A. Characteristics-Based Factor Model

Let r_t denote the stack of N excess returns. The conditional pricing equation is:

$$0 = \mathrm{E}_{t-1}\left[M_t \cdot r_t\right] . \tag{8}$$

One can find the SDF in the linear span of r_t (Hansen and Jagannathan, 1991):

$$M_t = 1 - a'_{t-1}\left(r_t - \mathrm{E}_{t-1}[r_t]\right) . \tag{9}$$

[4]Appendix B details the derivation.

This representation of SDF is time-contingent. To derive an unconditional pricing equation, characteristics-based asset pricing models assume that loadings on the return shocks vary proportionally to contemporaneous H firm characteristics (predictors):

$$a_{t-1} = Z_{t-1} b \tag{10}$$

where Z_{t-1} is a $N \times H$ predictor matrix. One can further isolate the time-invariant parameters (b) from predictor-weighted excess returns, known as factors (F_t):

$$M_t = 1 - b' Z_{t-1}' \left(r_t - \mathrm{E}_{t-1}[r_t]\right) \tag{11}$$

$$= 1 - b'(F_t - \mathrm{E}_{t-1}[F_t])\,. \tag{12}$$

Each entry in F_t is a linear combination of excess returns weighted by one predictor. Therefore, F_t is another collection of tradable excess returns, similar to r_t such that

$$\mathrm{E}[M_t \cdot F_t] = 0 \tag{13}$$

Even though Equation (9) and (12) are mathematically equivalent, the vast majority of asset pricing literature focuses on the unconditional representation of SDF. It not only condenses the cross section of myriad assets into characteristics-managed factors, but also does so with time-invariant loadings.

It is obvious from Equation (11) why the selection between levered and unlevered returns matters even if they span each other. Much of the existing literature backs their choice of factors with theories motivating why corresponding characteristics should predict returns. Hence the predictor matrix Z_{t-1} is economically informative. Employing a "wrong" set of returns in Equation (11) would require adjustments of Z_{t-1} to approximate "the true SDF". One can either introduce new characteristics to Z_{t-1} while keeping existing ones unaltered, or operating on existing characteristics in Z_{t-1} (*e.g.*, PCA) while keeping a lean dimension. The former brings about spurious factors, and the latter obscures economic interpretations. These issues are even more relevant given the non-relationship between levered and unlevered returns. Selecting a better set of returns can alleviate both drawbacks.

Solving the system of Equation (12) and (13) produces the coefficient b of the factor model:

$$b = \Sigma^{-1}\mathrm{E}\left[F_t\right] = \left(\Sigma\Sigma\right)^{-1}\Sigma\mathrm{E}\left[F_t\right] \tag{14}$$

where $\Sigma \equiv \mathrm{E}\left[\left(F_t - \mathrm{E}\left[F_t\right]\right)\left(F_t - \mathrm{E}\left[F_t\right]\right)'\right]$. b is the coefficients in a cross-sectional regression of the factors' population mean on its variance-covariance matrix. Empirically, regressions of the sample equivalents are used to estimate the coefficients. If we define:

$$\bar{\mu} = \frac{1}{T}\sum_{t=1}^{T} F_t \tag{15}$$

$$\bar{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\left(F_t - \bar{\mu}\right)\left(F_t - \bar{\mu}\right)', \tag{16}$$

the resulting estimator of b becomes:

$$\hat{b} = \bar{\Sigma}^{-1}\bar{\mu} = \left(\bar{\Sigma}\bar{\Sigma}\right)^{-1}\bar{\Sigma}\bar{\mu}\,. \tag{17}$$

B. Economically Motivated Bayesian Model by KNS

The risk of overfitting becomes substantial when a large number of candidate factors are considered. With the expansion of a factor model comes a higher propensity of picking up noises and performing poorly OOS. Regularziation methods are needed for model selection and mitigation of overfit. In the context of characteristics-based asset pricing model, the main source of overfit comes from the sample mean estimator $\hat{\mu}$ even with long samples, not from covariance (Kozak, Nagel, and Santosh, 2020). Following KNS, I proceed under the assumption that $\bar{\Sigma} = \Sigma$ and tackle with the imprecision of $\bar{\mu}$ by introducing a prior[5]:

$$\mu \sim \mathcal{N}\left(0, \frac{\kappa^2}{\tau}\Sigma^2\right) \tag{18}$$

where $\tau = \text{tr}\left[\Sigma\right]$ is the trace of the variance-covariance matrix and κ governs the strength of the prior. This prior implies an economically plausible notion that there exists a connection between the first and second moments of factor returns. Specifically, Sharpe ratios of factors associated with high-eigenvalue PCs should be higher than those associated with low-eigenvalue PCs[6]. It is statistically in line with many asset classes including stock returns that a few high-eigenvalue PCs account for most return variance while the contribution of the rest is neglegible.

With normal prior and likelihood, the Bayesian posterior mean and variance of b with a sample size of T are[7]:

$$\hat{b} = \left(\Sigma + \gamma I\right)^{-1} \bar{\mu} \tag{19}$$

$$\text{Var}(b) = \frac{1}{T}\left(\Sigma + \gamma I\right)^{-1} \tag{20}$$

where $\gamma = \frac{\tau}{\kappa^2 T}$. Compared with $\hat{b}$ in Equation (17), Equation (19) shrinks cofficients toward 0, similar to ridge regressions (Hastie et al., 2009). The effect is disproportionately stronger for factors associated with low-eigenvalue PCs. In simpler words, the KNS estimators are more intolerant toward coefficients of factors that "contribute" less to the cross-sectional variations. κ (or equivalently, γ) regularizes the fitting process in an economically plausible way. In fact, Equation (19) is the closed form solution to

[5] This family of priors is widely used in earlier literature. See Pástor (2000), Pástor and Stambaugh (2000), and Liechty, Harvey, and Liechty (2008).

[6] To see this, think of factors as orthogonized PCs and $\Sigma^{-1/2}\mu \sim \mathcal{N}\left[0, (\kappa^2/\tau)\Sigma\right]$.

[7] Inserting the Bayesian posteror $\hat{\mu} \sim \mathcal{N}\left[\left(\gamma I + \Sigma\right)^{-1}\bar{\mu}, \frac{1}{T}\left(\Sigma^{-1} + \gamma\Sigma^{-2}\right)^{-1}\right]$ into Equaition (17).

minimizing the HJ-distance (Hansen and Jagannathan, 1991) with L^2 penalty:

$$\hat{b} = \underset{b}{\operatorname{argmin}} \, (\hat{\mu} - \Sigma b)' \, \Sigma^{-1} \, (\hat{\mu} - \Sigma b) + \gamma b' b \,. \tag{21}$$

To implement the estimation, the value of κ is needed. Parameters like κ are called hyperparameters in machine learning language. They control the behavior of a fitting process and are not learned from data directly. Instead, they are tuned between sessions for better OOS performance.

I adopt a standard K-fold cross-validation (CV) method for hyperparameter tuning. (1) First, I divide the historical data into training and testing period, and further contiguously divide training data into K equal subsamples; (2) Then, for each possible κ, I compute $\hat{b}$ applying Equation (19) for $K-1$ of these subsamples and evaludate the OOS performance on the single withheld subsample via $R^2_{oos} = 1 - [\bar{\mu}_o - \bar{\Sigma}_o \hat{b}][\bar{\mu}_o - \bar{\Sigma}_o \hat{b}]' / [\bar{\mu}_o' \bar{\mu}_o]$ where subscript o indicates sample moment from the withheld subsample; (3) Next, I repeating this procedure K times, each time treating a different subsample as the OOS data. I average the R^2_{oos} across these K estimates; and (4) Finally, I choose the optimal hyperparameter κ that generates the highest average of R^2_{oos} and evaluate the real OOS performance on the testing period.

Throughout the CV process, I chose $K = 3$ following KNS as a compromise between the estimation error in $\hat{b}$ and $\bar{\Sigma}_o$. Based on the prior (Equation (18)), κ has a natural economic interpretation. It is the square root of the expected maximum squared Sharpe ratio:

$$\kappa = \left(\mathrm{E} \left[\mu \Sigma^{-1} \mu \right] \right)^{\frac{1}{2}} \,. \tag{22}$$

Two optimal κ's will be generated during the estimations of levered and unlevered factors. A higher optimal κ not only suggests a better-behaved data as there is less necessity for regularization (higher λ), but also represents a closer proximity to "the real SDF", which by construction, signals the highest Sharpe ratio.

So far, the Bayesian approach shrinks coefficients to almost but not exact zero, keeping all factors in the model. To reinforce my argument that spurious factors might arise from constructing SDF with levered returns, another penalty to filter out certain factors is desirable. I follow KNS and add an additional L^1 penalty on top of L^2:

$$\hat{b} = \underset{b}{\operatorname{argmin}} \left(\hat{\mu} - \Sigma b\right)' \Sigma^{-1} \left(\hat{\mu} - \Sigma b\right) + \gamma_2 b' b + \gamma_1 \sum_{i=1}^{H} |b_i| \,. \tag{23}$$

Due to the geometry of L^1 norm, Equation (23) accomplishes automatic factor selection without imposing that the SDF is necessarily sparse. In other words, the number of factors with non-zero coefficients is another hyperparameter that is optimized through the CV process and serves as an indicator of the comparison between levered and unlevered SDFs. I solve the optimization problem in Equation (23) using LARS-EN algorithm (Zou and Hastie, 2005)[8].

C. Hypotheses

Note that the Bayesian approach does not necessarily impose penalties because optimal hyperparameters are driven by OOS data. Therefore, I can employ the same approach on different sets of factors built from levered and unlevered return spaces and compare the data-revealed optimal hyperparameters that carry economic meaning. I will detail factor building in Section III.

Since I argue that some factors are introduced to capture the non-existent anomalies when levered factors do not faithfully reflect the economic theories, the same characteristics-based factors from unlevered space should not survive the shrinkage.

Hypothesis 1 *Cross-validated optimal number of factors with non-zero coefficients is*

[8]My LARS-EN code is built upon KNS.

smaller in unlevered SDF estimation than in levered SDF estimation under the Bayesian approach with L^1 and L^2 penalty.

If unlevering stock returns brings us closer to "the true SDF", implied MVE portfolio should yield higher Sharpe ratio in the unlevered space. In the simplest case, if unlevered market is a better proxy for the total wealth, it should have higher Sharpe ratio than stock market, and serves a better benchmark for CAPM test.

Hypothesis 2 *Unlevered SDF outperforms levered SDF OOS: 1. With optimal hyper-parameters (κ and number of factors with non-zero coefficients), mean-variance-efficient portfolio implied by unlevered model generates higher Sharpe ratio. 2. Alphas of various portfolios are smaller when tested against unlevered market.*

These hypotheses point to a sparser, better-performing SDF via unlevering.

III. Data and Unlevered Returns

I obtain daily stock returns from CRSP for all firms listed in the NYSE, AMEX, and NASDAQ and supplement the data with the three-month Treasury-bill rate from FRED as proxy for risk-free rate from which I calculate individual excess returns.

While there are at least hundreds of stock-level predictive signals in published research[9], I build upon the list of Green, Hand, and Zhang (2017) and Gu, Kelly, and Xiu (2020) after weighing feasibility and quality. I construct a large set of 100 firm-level characteristics based on the cross-section of stock returns literature[10]. Appendix C.1 lists the source to these characteristics. Data on firm equities, financial statements, and macroeconomic variables is retrieved from CRSP, Compustat, and Amit Goyal's personal

[9]Harvey, Liu, and Zhu (2016) studied 316 firm characteristics and common factors.

[10]I adapt the code from the Machine Learning Toolbox by Adrien d'Avernas, Martin Waibel, and Chunjie Wang.

website to build the characteristics. To obtain predictor matrix on a daily frequency, I forward fill quarterly or annual accounting-based characteristics.

I exclude in the data financial firms with SIC code between 6000 and 6999 and small-firms whose market caps are below 0.01% of the aggregate market[11]. My sample begins in January 1951 and ends in December 2022 (71 years) and includes 7422 firms that on average account for 74.4% of the total market value.

Finally, I employ the Merton model for unlevering. Unlevered returns are simply the growth rates of market value of firms, defined as the sum of market value of equities and debts. While equity values are readily available, data on market value of debt is inaccessible. Under the assumption that the total value of a firm follows GBM, the Merton model argues that the equity of the firm is a call option on the underlying value of the firm with a strike price equal to the face value of the firm's debt. Symbolically, the Merton model is similar to the BSM option pricing model:

$$E = V\mathcal{N}(d_1) - e^{-r_f T} B\mathcal{N}(d_2) \tag{24}$$

$$d_1 = \frac{\ln(V/B) + (r_f + 0.5\sigma_v^2)T}{\sigma_v\sqrt{T}} \tag{25}$$

$$d_2 = d_1 - \sigma_v\sqrt{T} \tag{26}$$

where E is the market value of equity, V is the market value of the firm, σ_v and σ_e are the volatilities of the aseets and equities, B is the face value of debt, T is debt's time-to-maturity, and r_f is the instantaneous risk-free rate. The model also implies that σ_e and

[11] The illiquidity from small stocks might contaminate the analyses. Financial firms that can sustain very high leverage might also drive the results. I include financial firms in robustness tests, in which the sample expands to 8004 unique stocks and accounts for 90.3% of the total market value on average. The results persist.

σ_v are related by[12]:

$$\sigma_e = \left(\frac{V}{E}\right) \mathcal{N}(d_1)\sigma_v \,. \tag{27}$$

The Merton model translates the volatility and market value of equity into those of firm's asset with Equation (24) and (27). All variables except σ_v and V are either known or can be estimated: (i) E is the product of the firm's shares outstanding and its current stock price; (ii) σ_e is measured by the annualized realized volatility of daily stock returns in each month; (iii) B is the sum of the firm's current liabilities and one half of its long-term liabilities; (iv) r_f is measured by the annualized return on three-month Treasury-bill rate; and (v) $T = 1$.[13]

Instead of solving this two-equation system directly, I implement the iterative procedure proposed by Crosbie and Bohn (2003) and Vassalou and Xing (2004)[14] to avoid a statistical challenge posed by acute movements of market leverage: (1) I guess an initial value of $\tilde{\sigma_v} = \sigma_e[E/(E + B)]$ and insert it into Equation (24) to infer the market value of each firm $\tilde{V}$ every day for the previous month; (2) I calculate the implied log return on assets each day and and use the returns series to generate new estimates $\tilde{\sigma_v}$; and (3) Iterate the steps until $\tilde{\sigma_v}$ converges so the absolute difference in adjacent $\tilde{\sigma_v}$'s is less than 10^{-3}.

Now that I obtain two panels of levered ($r_{e,t}$) and unlevered ($r_{a,t}$), the last step before implementing the estimation is to rank-normalize the predictor matrix Z_{t-1} such that each factor in F_t is a zero-investment long-short portfolio. Following KNS, for each predictor

[12] Under the GBM assumption, the equity value satisfies the time-series process: $dV = \mu_v V dt + \sigma_v V dz$ where dz is a standard Wiener process. It follows from Ito's Lemma and Equation (24) that $\sigma_e = \left(\frac{V}{E}\right) \frac{\partial E}{\partial V} \sigma_v$. In the BSM model, $\frac{\partial E}{\partial V} = \mathcal{N}(d_1)$.

[13] The estimations of T and B are also adopted by Chang, d'Avernas, and Eisfeldt (2021) and Gilchrist and Zakrajšek (2012). As a robustness check, I also apply another set of estimations following Bharath and Shumway (2008) in which B is the total liabilities and σ_e is measured by annualized realized volatility of daily stock returns in each year. The results remain consistent.

[14] Bharath and Shumway (2008); Gilchrist and Zakrajšek (2012); Chang, d'Avernas, and Eisfeldt (2021) also adopt this procedure.

at each time, I obtain the rank-transformed value as the ratio of a firm's rank in the predictor over the number of firms. Next, I normalize the value by first demeaning the rank-transformed predictor cross-sectionally then dividing the value by sum of absolute deviations from the mean of all firms. Along with the characteristics-based factors, I add an additional market factor to capture the level of risk premia.

I construct two panels of factors, $F_{e,t}$ and $F_{a,t}$, to test under the Bayesian approach discussed in Section II.

$$F_{e,t} = Z'_{t-1} \cdot r_{e,t} \tag{28}$$

$$F_{a,t} = Z'_{t-1} \cdot r_{a,t} \tag{29}$$

Factors do not necessarily need to be tradable. To build an OOS trading strategy, investors can refer to the covariance between test assets and non-tradable factors. However, tradable factors can directly translate into an MVE portfolio that can be easily evaluated OOS. Trading on $F_{a,t}$ is formidable as it requires access to the bonds as well as loans for all public firms on a daily frequency. Given the challenge, I build an additional set of tradable unlevered factors under the assumption that all debts are risk-free.

$$F_{\hat{a},t} = Z'_{t-1} \cdot [\mathbf{1}_N - L_{t-1}]' \cdot r_{e,t} \tag{30}$$

where L_t is an $N \times 1$ book leverage matrix defined as the ratio of book liabilities to the sum of book liabilities and market caps. This naïve approach assumes $r_{d,t} = \mathbf{0}_N$ at all times. Due to the cross-sectional heterogeneity of leverage, $F_{\hat{a},t}$ is nontrivial compared to $F_{e,t}$. The risk-free investments in the long leg differ from short leg thereby essentially altering the relative positions for levered stocks in the cross section: they are not proportional to Z_{t-1} anymore. Doshi et al. (2019) adopted the method in their primary analyses and found it sufficient to subsume value and volatility premium.

$F_{a,t}$ and $F_{\hat{a},t}$ stand for two extremes of market segmentation. While $F_{a,t}$ assumes no market segmentations between corporate equity and debt markets, $F_{\hat{a},t}$ assumes equity holders have no access to the debt market at all. In practice, investors especially institutional ones can build unlevered portfolios to certain extent. My analyses of $F_{\hat{a},t}$ and $F_{a,t}$ establish the boundaries for practical implications before I include bond in my analyses for future research.

As discussed in Section II, falsefully constructing SDFs with levered returns might have enticed asset pricers to either expand or transform the predictor matrix to capture the non-linear transformation from leverage. The former corresponds to new factors and the latter encompasses PCA. KNS found that applying a PC rotation on levered factors lead to a more sparse SDF under the Bayesian approach. While the dimensionality is tamed, transformations of economically motivated factors obscure the interpretations for why certain predictors have powers whilst others do not. I facilitate the comparision by considering the fourth set of factors: principal components of $F_{e,t}$:

$$P_{e,t} = Q_e' F_{e,t} \tag{31}$$

where Q_e is the matrix of eigenvectors of Σ_e.

Eventually, my sample translates to four sets of characteristics-based factors ($F_{a,t}$, $F_{e,t}$, $F_{\hat{a},t}$, and $P_{e,t}$) from February 1951[15] to December 2022 on a daily frequency, built from the same set of 100 predictors.

[15] $F_{a,t}$ and $F_{\hat{a},t}$ start from October 1970 when firms first reported quarterly long-term debts on Compustat.

IV. Empirical Results

A. Model Selections

Table I: Optimal Hyperparameters under Singular- and Dual-Penalty

This table collects CV-implied optimal hyperparameters for eight sets of factors built from the same 100 anomalies. They are levered F_e, unlevered F_a, tradable unlevered $F_{\hat{a}}$, PCs of levered P_e (Panel (a)), and their respective beta-neutral factors after orthogonizing against the market (Panel (b)). For L^2-only regularization, the hyperparameter is the root expected SR^2 (κ); For L^1-L^2 regularization, the hyperparameters include κ and the number of non-zero coefficients.

Panel (a): factors including market

		F_e	F_a	$F_{\hat{a}}$	P_e
L^2 Penalty	κ	0.69	0.77	0.45	0.69
L^1-L^2 Penalty	κ	0.75	3.16	0.64	0.75
	none-zero factors	88	5	42	27

Panel (b): beta-neutral factors

		$\tilde{F}_e$	$\tilde{F}_a$	$\tilde{F}_{\hat{a}}$	$\tilde{P}_e$
L^2 Penalty	κ	0.46	0.36	0.24	0.46
L^1-L^2 Penalty	κ	0.46	1.84	0.28	0.46
	none-zero factors	99	3	40	59

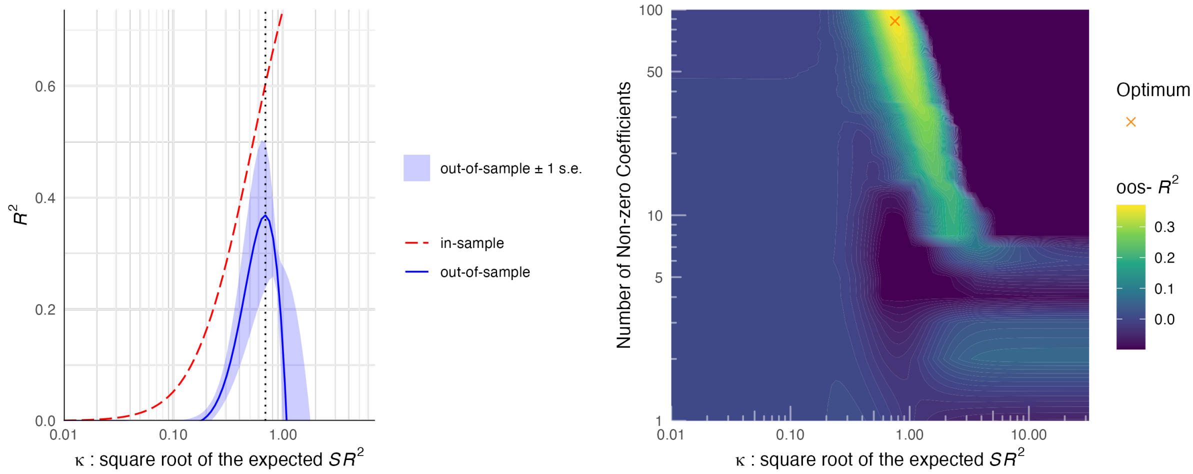
Figure 2 presents the OOS R^2 from the CV process for levered factors under both L^2-only and L^1-L^2 specifications with a range of hyperparameters.

When all factors are considered, the data calls for a sizable L^2-shrinkage to explain 36.8% of the OOS variation (Panel (a)). The in-sample (IS) R^2 decreases as I address the concerns of overfitting by imposing higher strength of the penalty in the Bayesian approach (lower λ and higher κ). As κ approaches 0, I use little IS information during coefficients estimation thereby IS R^2 converges to 0. On the contrary, relying too much (large κ) or too little (small κ) IS data leads bad OOS explanation. The model is optimized when κ is set around 0.75.

Similar OOS R^2 is only possible with the inclusion of most factors after allowing for sparsity (Panel (b)). The optimal number of factors with non-zero coefficients is 88, indicating little redundancy across factors. A small subset of these portfolios cannot span the SDF regardless of optimal κ. Forcing a sparse model would risk losing pricing information, shown as the significant drop in OOS R^2 moving down the plot. Despite my longer sample and wider collection of anomalies, the result is consistent with KNS that showed all 49 factors they built survived the shrinkage.

Figure 2: Levered Factors F_e: OOS R^2 under Singular- and Dual-Penalty

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using 100 anomaly portfolios of daily levered stock returns from 1951 to 2022. Panel (a) only employs L^2 penalty of which the strength is measured by prior root expected SR^2 (κ). Panel (b) also employs L^1 penalty of which the strength is measured by the number of retained factors. Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes of hyperparameters are ploted on logarithmic scale.



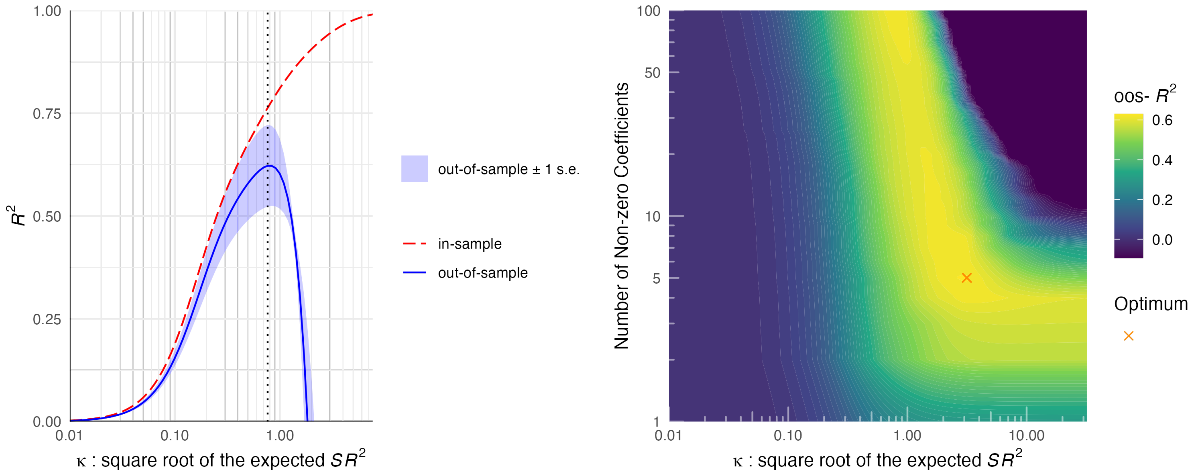
(a) L^2 penalty (b) L^1-L^2 penalties

Figure 3 presents the OOS R^2 under the same specifications and hyperparameters for unlevered factors. The situation is quite different: OOS R^2 is overall higher, indicating that unlevered factors in general carry more pricing information. While the optimal κ under singular penalty is comparable to its levered counterparty (0.77 versus 0.69), it is much higher under dual-penalty (3.16 versus 0.75), indicating that less supervision from L^2 penalty is needed when training unlevered factors: the concern of overfitting is milder and there is less noise in this space. Most importantly, It only requires 5 factors to peak the OOS R^2 at 62.4%: market, 12-month momentum, return on equity, asset growth, and revenue surprise. Compared with the optimized 88-factor levered SDF with 36.9% OOS R^2, unlevered data calls for a much sparser SDF and explains a much higher proportion of OOS variations. High R^2 in the contour plot covers a much wider area in the unlevered space than the levered one, implying some robustness in the vinicity around the optimum specification: it is inconsequential whether to include or exclude a few anomalies, or whether to have a slightly different root expected SR^2. In other words, additional factors provide little marginal benefit. In fact, they might introduce unnecessary noises and impair IS data's representativeness of the population, evidenced by the growing optimal L^2 penalty as more factors are included (the yellow strip runs from the southeast to the northwest in Figure 3 Panel (b)). Summing up, unlevering explains the cross-section better with fewer factors, suggesting some anomalies in the literature are due to the omission of non-linear transformation of returns.

As for the two complementary sets of factors, the optimal SDF implied by P_e and $F_{\hat{a}}$ are more sparse than F_e but less so than F_a. This is in line with the theory that asset pricers have been overstressing the predictor matrix Z while the choice of return space r is understudied. Even though model improvement can be achieved through optimizing Z (*e.g.* pricipal component or partial least squares), there is higher marginal gain by simply unlevering the return, even just under the wildest assumption that all debts are risk-free. The figures depicting their OOS R^2 against different hyperparameters can be found in Appendix D and key hyperparameters are collected in Table I.

Figure 3: Unevered Factors F_a: OOS R^2 under Singular- and Dual-Penalty

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using 100 anomaly portfolios of daily unlevered asset returns from 1970 to 2022. Panel (a) only employs L^2 penalty of which the strength is measured by prior root expected SR^2 (κ). Panel (b) also employs L^1 penalty of which the strength is measured by the number of retained factors. Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes of hyperparameters are ploted on logarithmic scale.

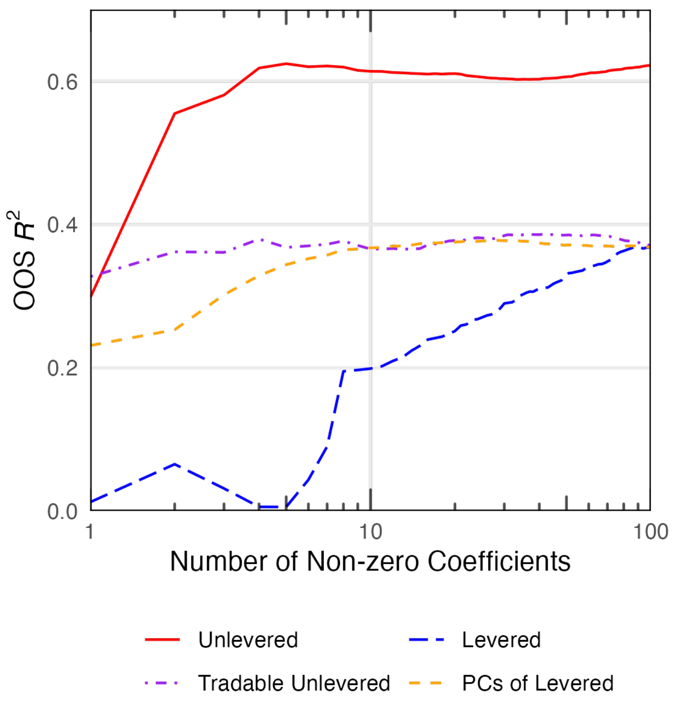


(a) L^2 penalty (b) L^1-L^2 penalties

To facilitate the comparison between factor sets, Figure 4 extracts a slice from the contour plots from Panel (b) of Figure 2 and 3 along the optimal κ for a given number of non-zero coefficients. I also include tradable factors $F_{\hat{a}}$ and PCs of levered factors P_e. As the figure shows, OOS R^2 only starts rising substantially for levered factors toward the right of the plot when more than 10 factors are admitted. This is consistent with KNS that showed it is never too much to add an additional factor proposed from the literature into the levered SDF. There is little redundancy in the return space. In constrast, a 5-factor model accounts for most variation in the unlevered space and very sparse models perform remarkably well. The marginal effect of adding an additional factor into the SDF is trivial after the 5-factor model and its OOS R^2 is much higher than that of the levered SDF. An optimized unlevered 5-factor SDF explains more variation than a levered 88-factor SDF optimized under the same approach. Tradable unlevered factors and PCs of levered factors are in between: sparse SDFs are sufficient to capture the cross-sectional variation but the ceiling of OOS R^2 is similar to the levered space. This result is striking: it provides evidence that even if the stock market and the debt market are entirely segmented and equity holders have zero access to the bond market, they can still benefit from reweighting the economically-motivated portfolios with a leverage matrix. Correctly accounting for firm-level, instead of stock-level risks, allows a small number of factor-mimicking portfolios that can compete against complicated synthetic portfolios built from conventional dimension reduction techniques. In reality, debt and stock markets are neither entirely segmented nor integrated. Households have limited access to the bond market, and institutions also to the loan market. Therefore, the unlevered and tradable unlevered cases outline the bounds of benefit from pricing firms' assets. It allows investors to construct an efficient portfolio from a mix of bonds and stocks managed by several economically founded predictors.

Figure 4: Entries of Predictors along Sparsity

This figure reports the maximum CV OOS R^2 under dual-penalty as specifications allow higher number of factors (x-axis). κ is optimized for respective number of non-zero coefficients.



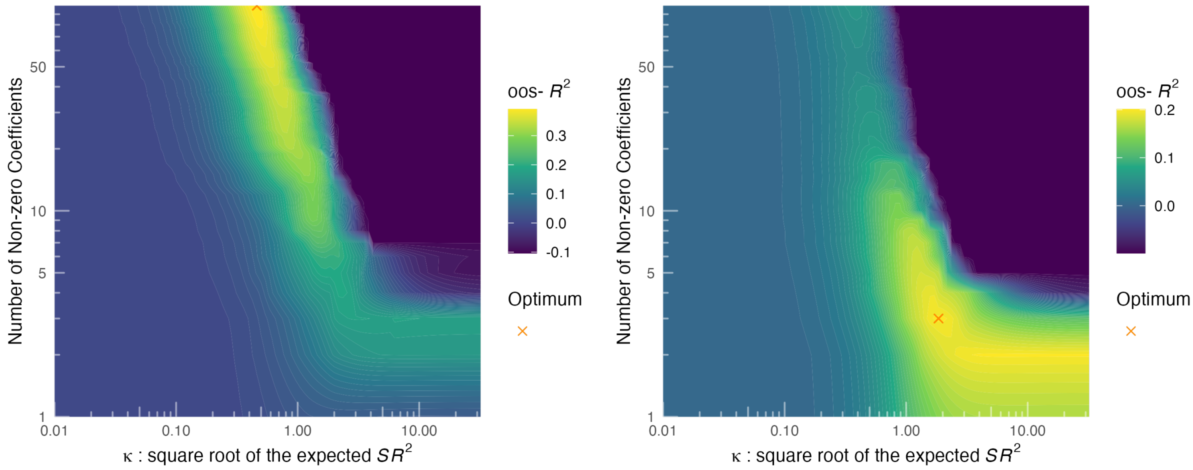
As shown in Table I, market survives all L^1 penalties regardlesss of the return spaces. It captures the level of equity or asset risk-premia. To focus on understanding the factors that help explain cross-sectional anomalies, I orthogonize every characteristics-based factors with respect to the market factor to examine the incremental power of other factors[16]. I denote these beta-neutral factors as $\tilde{F}_{a,t}$, $\tilde{F}_{e,t}$, $\tilde{F}_{\hat{a},t}$, and $\tilde{P}_{e,t}$ respectively and employ the exact same Bayesian methods.

[16]For each characteristics-based factor, I run a time-series regression on the market: $F_t \sim \beta\text{MKT}_t + \alpha$; then I calculate the corresponding beta-neutral factor as the remainder: $\tilde{F}_t = F_t - \hat{\beta}\text{MKT}_t$. For levered factor set F_e, the "market" refers to the value-weighted stock market returns; for unlevered set F_a, it refers to the value-weighted unlevered asset returns.

Figure 5 presents the OOS R^2 from the CV process for levered and unlevered beta-neutral factors under L^1-L^2 specifications with a range of hyperparameters. Similar to their counterparties with market, the optimal L^2 penalty (κ) is of the same order as 1. While the data still calls for the inclusion of most levered factors, the optimal specification in the unlevered space further reduces the dimensionality to three: 12-month momentum, return on equity, and revenue surprise. There are two important differences from F_a. First, the OOS R^2 of $\hat{F}_a$ nose dives across all specifications of hyperparameters after orthogonizing against market. This sharp drop is not observed in the levered space. In other words, including market in the factor model significantly enhances the OOS performance in the unlevered space, highlighting its contribution to explain the cross-sectional variation. Second, Panel (b) suggests that high OOS R^2 area clusters in the low dimension. In contrast, high OOS R^2 area covers both low and high dimensions when market is included: even though additional factors from F_a add little incremental benefit on top of the optimized 5-factor model, adding more factors does not impede the OOS prediction as long as a higher level of L^2 is imposed. On the contrary, additional factors from $F_{\hat{a}}$ on top of the optimized 3-factor model negative impact in OOS prediction in the beta-neutral case. The two differences are not observed in the levered space after orthogonizing against the market. These results suggest that there is much less left to explain on top of the market in the unlevered space, and many factors are mostly composed of noises after removing the market element from the factors because including them IS exacerbates the over-fitting problem. Figure D.3 in Appendix D reports the OOS R^2 of the two beta-neutral complementary sets of factors $\tilde{F}_{\hat{a}}$ and $\tilde{P}_e$.

Figure 5: OOS R^2 under Dual-Penalty for Beta-Neutral Factors

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using 100 anomaly portfolios. Panel (a) depicts the result for beta-neutral levered factors ($\tilde{F}_{e,t}$) and Panel (b) for beta-neutral unlevered factors ($\tilde{F}_{a,t}$). Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes are ploted on logarithmic scale.



(a) levered $\tilde{F}_e$ (b) unlevered $\tilde{F}_a$

Similar to Figure 4, Figure 6 takes a cut in the contour plots of Figure 5 along the ridge of maximal OOS R^2 from bottom to top where we optimize L^2 shrinkage (κ) for each level of sparsity. I also include beta-neutral tradable factors $\tilde{F}_{\hat{a}}$ and PCs of beta-neutral levered factors $\tilde{P}_e$. The trends persist compared to pre-orthogonizing cases across all four sets of factors except for the unlevered one, where we see a decline after peaking at 3-factor model. On the contrary, OOS R^2 keeps rising toward the right of the plot in the levered space, that is, adding additional anomalies to existing models is always marginally benefitial. In other three spaces, OOS R^2 already peak when less than a dozen factors are considered. Red (unlevered) and blue lines (levered) almost move in the opposite directions: as data learns more levered factors, more unlevered data only adds noises to the training. Again, beta-neutral unlevered factors have very limited explanatory power for cross-sectional variations not only compared to beta-neutral levered factors, but also to pre-orthogonizing unlevered factors. These results point to a CAPM-like parsimonious factor model in the unlevered space whilst levered factor models leave many anomalies unanswered.

Figure 6: Entries of Predictors along Sparsity (Beta-Neutral)

This figure reports the maximum CV OOS R^2 under dual-penalty as specifications allow higher number of beta-neutral factors (x-axis). κ is optimized for respective number of non-zero coefficients.

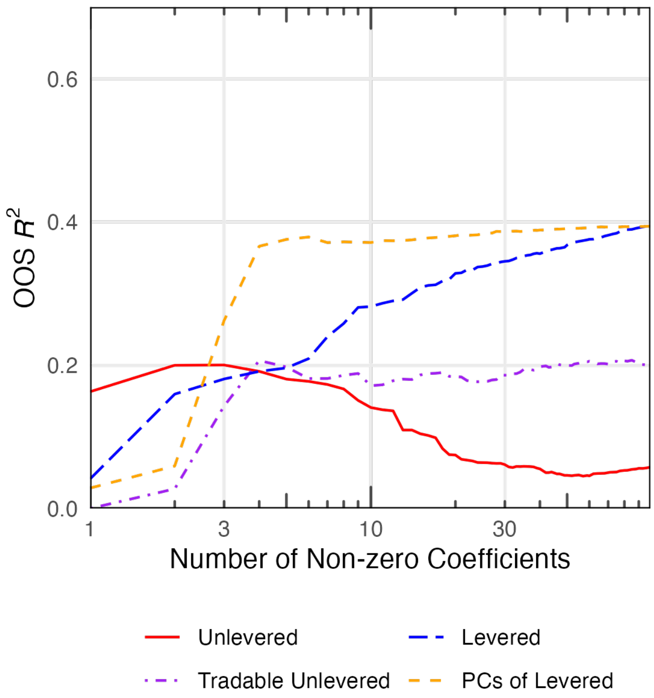


Table I collects optimal hyperparameters for all the cases discussed, based on which I build SDF-implied MVE portfolios to test later for their OOS performance. It is clear that unlevering significantly reduce the dimensionality required to capture the cross-sectional variation. On the other hand, the optimal L^2 regularizations are on similar levels for all return spaces, regardless of orthogonizing againt the market or not. In other words, there is evidence that the expansion of "the factor zoo" might have partly arisen from ignoring the non-linearity of the leverage effect, but the noises that challenge μ's evaluation persist thus similar level of regularization is still necessary to deal with overfitting.

B. Factor Importance

Equation (20) gives a closed-form solution to the posteiror standard error for the coefficient estimates under the singular penalty of κ. I report top 10 most significant factors in Table II. Market, momentums and ROE are robust across all factor sets. The t-statistics are low for most anomalies, but what is important is the joint significance of these factors and the explanatory power of the SDF constructed from them.

Table II: Coefficient Estimates under L^2 Penalty

Under a singular regularization of root expected SR^2 (κ), this table lists top 10 of 100 coefficient estimates and t-statistics corresponding to the CV-implied optimal prior, sorted by the absolute values of the t-statistics. Panel (a) reports two key sets of levered and Merton-unlevered factors. Panel (b) reports two complementary sets of factors: riskless-unlevered factors and PCs of levered factors.

Panel (a)

levered anomaly portfolios (F_e)			unlevered anomaly portfolios (F_a)		
predictors	b	t-stat	predictors	b	t-stat
market	4.181	4.458	12-month momentum	4.737	2.696
12-month momentum	3.551	2.654	revenue surprise	4.283	2.260
1-month momentum	-3.211	-2.451	return on equity	3.408	1.797
change in 6-month momentum	-2.915	-2.129	market	2.013	1.774
change in shares outstanding	-2.254	-1.521	6-month momentum	2.499	1.416
earnings to price	1.907	1.319	size (industry-adjusted)	-2.395	-1.260
R&D to sales	1.646	1.190	change in employees (industry-adjusted)	2.209	1.156
return on equity	1.626	1.142	volatility of liquidity (share turnover)	2.193	1.151
maximum daily return	-1.643	-1.123	change in 6-month momentum	-1.978	-1.107
number of earnings increase	1.594	1.065	number of earnings increase	2.088	1.082

Panel (b)

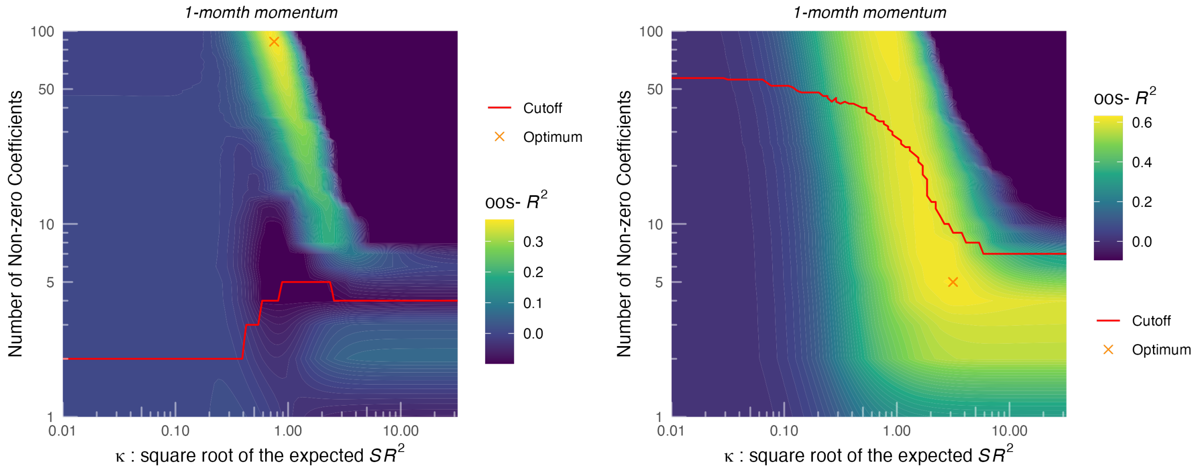
tradable unlevered anomaly portfolios ($F_{\hat{a}}$)			PCs of levered anomaly portfolios (P_e)		
predictors	b	t-stat	predictors	b	t-stat
market	1.871	2.321	PC3	4.973	5.213
12-month momentum	1.498	1.388	PC10	5.710	4.537
1-month momentum	-1.091	-1.004	PC2	-1.530	-2.121
R&D to market capitalization	1.025	0.920	PC5	-2.224	-2.080
change in 6-month momentum	-0.969	-0.882	PC12	2.361	1.814
return on assets	0.929	0.829	PC6	-1.740	-1.551
return on equity	0.886	0.787	PC25	1.963	1.374
change in shares outstanding	-0.866	-0.766	PC28	-1.829	-1.271
return on invested capital	0.767	0.682	PC9	-1.550	-1.242
financial statement score	0.767	0.681	PC14	-1.618	-1.224

A rank of significance for factors when allowing for sparsity is more informative because it hints anomalies that might have been introduced to capture the non-linearity between the levered and unlevered returns, rather than reflect the fundamental macroeconomic risk that affect the firm. Without a closed-form solution to the standard error of estimates under dual-penalty similar to Equation (20), I rank the factor importance by their earliest entry into the the SDF when allowing higher dimensions (moving from bottom to top in the contour plots in Figure 2 and 3). Comparing components in levered and unlevered sparse SDFs of the same length, factors only appearing in the levered SDF are either capturing economically-founded risks that are specific to the stock market, or the non-linear return transformation from the leverage.

For example, Figure 7 shows the cutoffs above which the coefficient estimate of 1-month momentum factor is non-zero and below which is zero. The lower the line is, the more important the corresponding factor is. In this specific case, 1-month momentum factor shall be included in all levered SDFs as long as more than 5 factors are allowed while it would only be selected in unlevered SDFs with more than 20 factors on average. In other words, unlevering drags 1-month momentum down the priority list. The rankings can supplement the optimal coefficients estimates when comparing the F_a against F_e regardless of whether the factor in question is non-zero at optimums: whether to have a 5- or 6-factor model does not make a huge difference, especially when different sample might hint slightly different optimums, but a major decrease in such ranking after unlevering might imply close link to the leverage.

Figure 7: Factor Selections: an Example

This figure demonstrates the cutoffs of 1-month momentum on top of the contour map of OOS R^2 under dual penalty for levered (Panel (a)) and unlevered (Panel (b)) factors. The coefficient estimate for the specific factor is zero below the line and non-zero above the line.



(a) Levered F_e (b) Unlevered F_a

To quantify the shift in the cutoff, I rank each characteristic by its first entry into the SDF when relaxing the number of admitted factors while setting κ corresponding to the particular sparsity at optimum. Table III demonstrates top 10 factors for each return spaces. For instance, most momentum-based factors appear in a sparse stock pricing model but all are trivial after unlevering except 12-month momentum. On the other hand, there is evidence that labor and profitability measures are of much higher ranks among unlevered factors. Remember that empirical data requires 88 levered factors to summarize the cross section, thus the result does not suggest labor and profitability are not or less important for stock pricing.

Even though it is out of the scope of this paper to inspect the economic interpretation behind how specific factors might have unintendedly capture the leverage effect or statistical errors accumulated from previous ill-defined factors, rather, I aim to measure the joint severity of the issue, this exercise still sheds a light on what anomalies need further investigation on why they diverge in their ranks in asset versus stock pricing.

Table III: Factor Importance

This table reports top 10 factors ranked by the sequence of entry into the SDF when allowing for higher number of non-zero coefficients under dual-penalty. Factors are selected to generate highest OOS R^2 from 3-fold cross validation process with 100 candidate anomaly portfolios. The other hyperparameter root expected SR^2 (κ) is set at respective optimum for each number of non-zero coefficients.

earliest entry	levered (F_e)	unlevered (F_a)	tradable unlevered ($F_{\hat{a}}$)
1	market	market	market
2	12-month momentum	12-month momentum	12-month momentum
3	1-month momentum	revenue surprise	change in shares outstanding
4	6-month momentum	return on equity	1-month momentum
5	change in shares outstanding	asset growth	maximum daily return
6	sales to price	employee growth rate	earnings to price
7	industry momentum	change in employees (industry-adjusted)	asset growth
8	change in 6-month momentum	earnings volatility	return volatility
9	earnings to price	maximum daily return	change in 6-month momentum
10	maximum daily return	size (industry-adjusted)	industry momentum

C. OOS Performance of Market and SDF-implied MVE Portfolios

The discussion so far is restricted to hyperparameter tuning. Once κ and number of factors are set to the optimums, I can proceed to build MVE portfolios and compare their OOS performance. I re-estimate $\hat{b}$ under the optimal hyperparameter excluding a time window as testing period. Following KNS, I set the testing period from January 2005 to December 2022[17]. The OOS MVE portfolo for each factor set is given by:

$$MVE_t = \hat{b}' \cdot F_t \,. \tag{32}$$

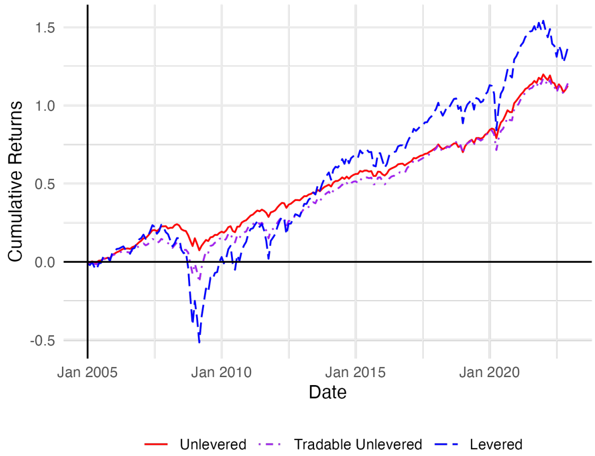
[17] In robustness tests, I also set different start dates for the testing period: January 2000 and January 2010. The results persist.

Figure 8 and Table IV report the cumulative log returns and the Sharpe ratios for markets and SDF-implied MVE portfolios. Among the market returns, stock market has the highest mean and volatility. This is no surprise because market portfolios long risky securities and short risk-free rate. Equities as call options to firms assets, are exposed to higher systematic risks.

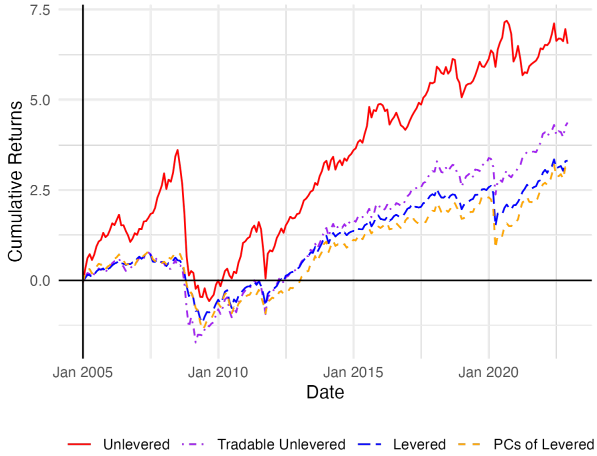
Under classic asset pricing theory, MVE frontier is an equivalent representation of the SDF thereby carrying the highest Sharpe ratio, thus I focus on comparing the Sharpe ratio of key portfolios. In an bootstrapping exercise, I randomly resample returns of the same length with replacement from all spaces and calculate their respective Sharpe ratio. Repeating the step for 1000 times yields empirical standard errors, which I report in the parentheses in Table IV. Levered stock market is less efficient than unlevered asset market, with lower Sharpe ratio (0.46 versus 0.98). On the other hand, after we add efficient factor-mimicking portfolios to the markets, the cumulative returns all drastically increase, with unlevered efficient portfolio leading the rest. Interestingly, the Sharpe ratio of unlevered efficient portfolio is not significantly different from unlevered market portfolio (0.98 versus 0.89), indicating that the added 4 portfolios are not closer to "the true SDF" than the market. In constrast, adding 87 addition factors would lift the Sharpe ratio from stock market closer to what unlevered SDF implies and their difference is less siginificant under the same bootstrapping method. However, on average levered SDF consisting of 87 factors still underperforms compared to 5 unlevered factors that were optimized under the same Bayesian methods (0.63 versus 0.89).

Figure 8: Cumulative Return of Markets SDFs-implied MVE portfolios

This figure demonstrates the OOS daily cumulative logarithmic returns of Market (Panel (a)) and SDF-implied MVE (Panel (b)) portfolios from 2005 to 2022 implied by four sets of SDFs in levered, unlevered, tradable unlevered, and PCs of levered spaces. Model specifications (κ and number of factors) are trained and optimized using data from 1951 to 2004.



(a) market portfolios



(b) MVE portfolios

Table IV: OOS Cumulative Returns and Sharpe Ratios of MVEs and MKTs: 2005-2022

This table reports the cumulative logarithmic returns and Sharpe ratios of six portfolios: market and levered SDF-implied MVE portfolios from levered, tradable unlevered and unlevered spaced. MVE portfolios are optimized with sample from Feb 1951 to December 2004. The time priod is from January 2005 to December 2022. Empirical standard errors are derived from bootstrapping with 1000 resamples and reported in the parentheses.

	market portfolios			implied MVE portfolios		
	levered	tradable unlevered	unlevered	levered	tradable unlevered	unlevered
log cumulative return	1.31	1.11	1.10	3.30	4.36	6.60
Sharpe ratio	0.46	0.68	0.98	0.63	0.71	0.89
bootstrap s.e.	(0.24)	(0.24)	(0.24)	(0.23)	(0.24)	(0.25)

In the previous subsection, I briefly touched upon how unleverd space might bring us closer to CAPM. Now I formally inspect the "abnormal returns" of MVE protfolios against various market benchmarks: the stock market return, Fama-French 4 factors returns, the tradable unlevered asset market return, and the unlevered market return. If my hypotheses are correct that many anomalies arise from constructing the SDF from a misspecified levered return space, then assets' α's should drop in both economic scale and statistcal significance along these benchmarkes that employ progressively better return space. I choose the MVE portfolios as testing assets to further investigate their performance. The results are indeed consistent with my hypotheses.

Table V reports the annualized abnormal returns α's (in %) from time-series regressions of MVE portfolios on benchmark portfolios. Overall, the market portfolios vary from least to most plausible (from top to bottom), as suggested by the drop of magnitude and significance of α for all three testing assets; In addition, the testing assets enjoy gradually higher abnormal returns (from left to right), as suggested by the increase of magnitude and significance of α for all benchmark portfolios. To be more specific, I unsurprisingly uncovered the anomalies studied extensively in stock pricing literature by regressing a 88-factor model on market (1st row, 1st column). However, once you correctly construct a new market return in the unlevered space, the anomaly vanishes (3rd and 4th row, 1st column), thereby CAPM explains away many anomalies proposed in the literature. There still exists unexplained returns: after all, this improved market return cannot explain all the variations of the unleverd SDF implied MVE portfolios (last row, last column). The good news is, I shrink the 87 cross-sectional anomalies into 4; the bad news is: the unlevered alpha is material. Nevertheless, we are much closer to a parsimonious CAPM specification after unlevering.

Table V: Annualized α of MVE Portfolios against Various Benchmarks

This table measures the distance to CAPM by checking the annualized α (in %) from regressing SDF-implied MVE portfolios on various market portfolios and Fama-French 4 factors (market, size, value, and profitability). The SDFs are optimized on dual-penalty. MVE portfolio and benchmark returns are normalized to have the same standard deviation as the aggregate stock market for better comparison. Standard errors are reported in parentheses.

benchmarks \test assets	levered MVE	tradable unlevered MVE	unlevered MVE
stock market	5.50^{*}	6.43^{**}	12.76^{***}
	(3.03)	(2.53)	(3.95)
Fama-French 4 Factors	3.53^{*}	4.99^{**}	9.65^{***}
	(2.03)	(2.04)	(2.70)
tradable unlevered market	2.31	3.11	11.16^{***}
	(3.13)	(2.76)	(4.14)
unlevered market	−1.92	−1.48	7.88^{***}
	(3.20)	(2.86)	(4.10)
Note:			*p<0.1; **p<0.05; ***p<0.01

V. Conclusion

Most economically founded stock pricing factors reveal risk profile on a firm level. Building factor-mimicking portfolios in the stock return space does not faithfully reflect the fundamental macroeconomic risks that are proxies for the utility growth. Equities are call options to the firms' assets and the non-linear return structures can distort the risk drivers suggested by theoretical works. The accumulated statistical errors might have in part responsible to the neverending expansion of "the factor zoo".

In this paper, I shrink the cross-sectional variations of levered stocks and Merton-unlevered assets with the same set of 100 firm return predictors proposed by the literature. I employ an economically-motivated Bayesian prior following Kozak, Nagel, and Santosh (2020) to regularize the high dimensions. Cross validated 5-factor unlevered SDF outperforms its 88-factor levered counterparty in sparsity, R^2, OOS Sharpe ratio, as well as market alphas. CAPM is more supported in the unlevered space, evidenced by a much lower explanatory power of unlevered factors after orthogonizing against the market, and a lack of improvement in the Sharpe ratio on top of asset market.

These results attribute a substantial number of stock pricing anomalies to the failure of projecting the SDF onto the asset return space that truthfully reflects firm risks. It is much more economically coherent to price assets, not stocks.

References

Bali, T. G., H. Beckmeyer, et al. (2023). "Option return predictability with machine learning and big data". In: *The Review of Financial Studies* 36(9), pp. 3548–3602.

Berk, J. B. and P. M. DeMarzo (2007). *Corporate finance*. Pearson Education.

Berk, J. B., R. C. Green, and V. Naik (1999). "Optimal investment, growth options, and security returns". In: *The Journal of finance* 54(5), pp. 1553–1607.

Bharath, S. T. and T. Shumway (2008). "Forecasting default with the Merton distance to default model". In: *The Review of Financial Studies* 21(3), pp. 1339–1369.

Carlson, M., A. Fisher, and R. Giammarino (2004). "Corporate investment and asset price dynamics: Implications for the cross-section of returns". In: *The Journal of Finance* 59(6), pp. 2577–2603.

Chang, H., A. d'Avernas, and A. L. Eisfeldt (2021). "Bonds vs. equities: Information for investment". In: *Equities: Information for Investment (April 15, 2021)*.

Christoffersen, P., S. Heston, and K. Jacobs (2013). "Capturing option anomalies with a variance-dependent pricing kernel". In: *The Review of Financial Studies* 26(8), pp. 1963–2006.

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

Coval, J. D. and T. Shumway (2001). "Expected option returns". In: *The journal of Finance* 56(3), pp. 983–1009.

Crosbie, P. and J. Bohn (2003). "Modeling default risk: Modeling methodology". In: *KMV corporation*.

Doshi, H., K. Jacobs, et al. (2019). "Leverage and the cross-section of equity returns". In: *The Journal of Finance* 74(3), pp. 1431–1471.

Fama, E. F. (1991). "Efficient capital markets: II". In: *The journal of finance* 46(5), pp. 1575–1617.

Fama, E. F. and J. D. MacBeth (1973). “Risk, return, and equilibrium: Empirical tests”. In: *Journal of political economy* 81(3), pp. 607–636.

Feng, G., S. Giglio, and D. Xiu (2020). “Taming the factor zoo: A test of new factors”. In: *The Journal of Finance* 75(3), pp. 1327–1370.

Gilchrist, S. and E. Zakrajšek (2012). “Credit spreads and business cycle fluctuations”. In: *American economic review* 102(4), pp. 1692–1720.

Green, J., J. R. Hand, and X. F. Zhang (2017). “The characteristics that provide independent information about average US monthly stock returns”. In: *The Review of Financial Studies* 30(12), pp. 4389–4436.

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

Hansen, L. P. and R. Jagannathan (1991). “Implications of security market data for models of dynamic economies”. In: *Journal of political economy* 99(2), pp. 225–262.

Harvey, C. R., Y. Liu, and H. Zhu (2016). “… and the cross-section of expected returns”. In: *The Review of Financial Studies* 29(1), pp. 5–68.

Hastie, T., R. Tibshirani, et al. (2009). *The elements of statistical learning: data mining, inference, and prediction.* Vol. 2. Springer.

Kozak, S., S. Nagel, and S. Santosh (2020). “Shrinking the cross-section”. In: *Journal of Financial Economics* 135(2), pp. 271–292.

Liechty, J., C. R. Harvey, and M. W. Liechty (2008). “Bayes vs. resampling: A rematch”. In: *Journal Of Investment Management* 6(1).

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

Pástor, L. (2000). “Portfolio selection and asset pricing models”. In: *The Journal of Finance* 55(1), pp. 179–223.

Pástor, L. and R. F. Stambaugh (2000). "Comparing asset pricing models: an investment perspective". In: *Journal of Financial Economics* 56(3), pp. 335–381.

Schlag, C. and T. Sichert (2020). "The shape of the pricing kernel and expected option returns". In: *Swedish House of Finance Research Paper*, pp. 20–25.

Vassalou, M. and Y. Xing (2004). "Default risk in equity returns". In: *The journal of finance* 59(2), pp. 831–868.

Zou, H. and T. Hastie (2005). "Regularization and variable selection via the elastic net". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 67(2), pp. 301–320.

.

Appendix A. Expected Excess Stock Return and Debt

This appendix details the derivation of the relation between expected excess stock return and the face value of debt of a firm following Coval and Shumway (2001).

Assuming the existence of an SDF that prices all assets with:

$$1 = \mathrm{E}\left[M \cdot R\right] \tag{A.1}$$

where R is the gross return of any asset, and M is the strictly positive SDF. Denote the firm's face value of debt as B and firm's market value of assets on maturity date as v that is a random varible followsing probability density distribution $f(v)$. The expected excess equity return is

$$\mathrm{E}[r_e(B)] = \frac{\mathrm{E}[\max(v-B,0)]}{\mathrm{E}[M \cdot \max(v-B,0)]} - 1 \tag{A.2}$$

$$= \frac{\int_{v=B} (v-B) f(v) \partial v}{\int_{m=0} \int_{v=B} m(v-B) f(v,m) \partial v \partial m} - 1 \tag{A.3}$$

$$= \frac{\int_{v=B} (v-B) \left[1 - \mathrm{E}[M|v]\right] f(v) \partial v}{\int_{v=B} (v-B) \mathrm{E}[M|v] f(v) \partial v} \tag{A.4}$$

where $f(v,m)$ is the joint distribution of the asset value and the SDF. Applying Leibniz inegral rule, the derivative of expected net returns with respect to the debt can be expressed as

$$\frac{\partial \mathrm{E}[r_e(B)]}{\partial B} = \frac{\int_{v=B}(v-B)f(v)\partial v \cdot \int_{v=B} \mathrm{E}[M|v]f(v)\partial v - \int_{v=B}(v-B)\mathrm{E}[M|v]f(v)\partial v \cdot \int_{v=B} f(v)\partial v}{\int_{v=B}(v-B)\mathrm{E}[M|v]f(v)\partial(s)} \tag{A.5}$$

$$= \frac{\int_{v=B} \frac{v-B}{1-F(B)} f(v)\partial v \cdot \int_{v=B} \frac{\mathrm{E}[M|v]}{1-F(B)} f(v)\partial v - \int_{v=B} \frac{(v-B)\mathrm{E}[M|v]}{1-F(B)} f(v)\partial v}{\left[\int_{v=B}(v-B)\,\mathrm{E}\left[M|v\right] \frac{f(v)}{1-F(B)}\partial v\right]^2} \tag{A.6}$$

where $F(v)$ is the corresponding cumulative density for $f(v)$. The numerator and denominator of Equation (A.6) are composed of several conditional expectations that can be rewritten as

$$\frac{\mathrm{E}[M|v>B] \cdot \mathrm{E}[v-B|v>B] - \mathrm{E}[\mathrm{E}(M|v)(v-B)|v>B]}{(\mathrm{E}[\mathrm{E}(M|v)(v-B)|v>B])^2} \tag{A.7}$$

that can be further simplified as

$$\frac{-\mathrm{Cov}\left[\mathrm{E}\left(M|v\right), v-B|v>B\right]}{(\mathrm{E}[\mathrm{E}(M|v)(v-B)|v>B])^2}\,. \tag{A.8}$$

When the SDF moves against the underlying firm's market value of assets conditional on the firm being solvent, the derivative is positive.

Appendix B. Characteristics-Based Factor Model

I postulate a stochastic discount factor (SDF) that is projected to the payoff space $\underline{X}$ to be a linear function of the shocks to the payoffs (Hansen and Jagannathan, 1991):

$$x^* = \mathrm{E}[x^*] + (x - \mathrm{E}[x])' a \,. \tag{B.1}$$

For any arbitrary asset whose payoff is x, its price p satisfies

$$\begin{aligned} p &= \mathrm{E}[x x^*] && \text{(B.2)} \\ &= \mathrm{E}[x^*]\mathrm{E}[x] + \mathrm{E}[x(x - \mathrm{E}[x]]')a && \text{(B.3)} \\ &= \mathrm{E}[x^*]\mathrm{E}[x] + \mathrm{E}\left[(x - \mathrm{E}[x])(x - \mathrm{E}[x])'\right] a && \text{(B.4)} \\ &= \mathrm{E}[x^*]\mathrm{E}[x] + \Omega a \,. && \text{(B.5)} \end{aligned}$$

We can solve for a and insert it back to Equation (B.1):

$$\begin{aligned} a &= \Omega^{-1} \left(p - \mathrm{E}[x^*]E[x]\right) && \text{(B.6)} \\ x^* &= \mathrm{E}[x^*] + \left(p - \mathrm{E}[x^*]\mathrm{E}[x]\right)' \Omega^{-1} \left(x - \mathrm{E}[x]\right) \,. && \text{(B.7)} \end{aligned}$$

Considering excess returns in the payoff space and pick a random zero-beta rate $R_f = 1$ gives

$$\begin{aligned} p &= 0 && \text{(B.8)} \\ x &= r && \text{(B.9)} \\ E(x^*) &= \frac{1}{R_f} = 1 \,. && \text{(B.10)} \end{aligned}$$

Denote the SDF as M, then Equation (B.7) becomes

$$M = 1 - \mathrm{E}[r]\Omega^{-1}\left(r - \mathrm{E}[r]\right) . \tag{B.11}$$

In a multiperiod world

$$M_t = 1 - a'_{t-1}\left(r_t - \mathrm{E}_{t-1}[r_t]\right) \tag{B.12}$$

where a_{t-1} is the product of two expectations($\mathrm{E}_{t-1}[r]$ and $\mathrm{E}_{t-1}^{-1}\left[(r - \mathrm{E}[r])(r - \mathrm{E}[r])'\right]$), known at $t-1$ before r_t are realized at t. Characteristics-based asset pricing models assume the loadings on the return shocks are linear combinations of return predictors (*e.g.* firm characteristics and macroeconomic variables) and parametrize a_{t-1} as

$$a_{t-1} = Z_{t-1}b \tag{B.13}$$

where Z is a $N \times H$ predictors matrix. N is the number of assets in the economy and H is the number of predictors. To clarify the decomposition, I assume only two predictors $i_{n,t}$ and $j_{n,t}$ in the economy where the subscripts represent the cross section and time series respectively:

$$a_{t-1} = \begin{bmatrix} b_i i_{1,t-1} + b_j j_{1,t-1} \\ b_i i_{2,t-1} + b_j j_{2,t-1} \\ \dots \\ b_i i_{N,t-1} + b_j j_{N,t-1} \end{bmatrix} = \begin{bmatrix} i_{1,t-1} & j_{1,t-1} \\ i_{2,t-1} & j_{2,t-1} \\ \dots & \\ i_{N,t-1} & j_{N,t-1} \end{bmatrix} \cdot \begin{bmatrix} b_i \\ b_j \end{bmatrix} = Z_{t-1}b\,. \tag{B.14}$$

Inserting Equation (B.13) back to Equation (B.12), gives

$$M_t = 1 - b' Z'_{t-1} \left(r_t - \mathrm{E}_{t-1}[r_t] \right) \tag{B.15}$$

$$= 1 - \begin{bmatrix} b_i & b_j \end{bmatrix} \cdot \begin{bmatrix} i_{1,t-1} & i_{2,t-1} & \cdots & i_{N,t-1} \\ j_{1,t-1} & j_{2,t-1} & \cdots & j_{N,t-1} \end{bmatrix} \begin{bmatrix} r_{1,t} - \mathrm{E}_{t-1}[r_{1,t}] \\ r_{2,t} - \mathrm{E}_{t-1}[r_{2,t}] \\ \cdots \\ r_{N,t} - \mathrm{E}_{t-1}[r_{N,t}] \end{bmatrix} \tag{B.16}$$

$$= 1 - b' \begin{bmatrix} i_{1,t-1}(r_{1,t} - \mathrm{E}_{t-1}[r_{1,t}]) + \cdots + i_{N,t-1}(r_{N,t} - \mathrm{E}_{t-1}[r_{N,t}]) \\ j_{1,t-1}(r_{1,t} - \mathrm{E}_{t-1}[r_{1,t}]) + \cdots + j_{N,t-1}(r_{N,t} - \mathrm{E}_{t-1}[r_{N,t}]) \end{bmatrix} \tag{B.17}$$

$$= 1 - b' \left(\begin{bmatrix} i_{1,t-1} r_{1,t} + \cdots + i_{N,t-1} r_{N,t} \\ j_{1,t-1} r_{1,t} + \cdots + j_{N,t-1} r_{N,t} \end{bmatrix} - \begin{bmatrix} i_{1,t-1} \mathrm{E}_{t-1}[r_{1,t}] + \cdots + i_{N,t-1} \mathrm{E}_{t-1}[r_{N,t}] \\ j_{1,t-1} \mathrm{E}_{t-1}[r_{1,t}] + \cdots + j_{N,t-1} \mathrm{E}_{t-1}[r_{N,t}] \end{bmatrix} \right) \tag{B.18}$$

$$= 1 - b'(F_t - \mathrm{E}_{t-1}[F_t]) \,. \tag{B.19}$$

Each element in F_t is a linear combination of excess returns weighted by one predictor, thus also tradable such that

$$\mathrm{E}[M_t \cdot F_t] = 0 \,. \tag{B.20}$$

Solving the system of Equation (B.19) and (B.20) results in the coefficient b of the factor model:

$$b = \Sigma^{-1} \mathrm{E}\left[F_t\right] = \left(\Sigma\Sigma\right)^{-1} \Sigma \mathrm{E}\left[F_t\right] \tag{B.21}$$

where $\Sigma \equiv \mathrm{E}\left[\left(F_t - \mathrm{E}\left[F_t\right]\right)\left(F_t - \mathrm{E}\left[F_t\right]\right)'\right]$. Empirically, b is the coefficients in a cross-sectional regression of the factors' population mean on its variance-covariance matrix.

In a special case, demeaning Z_{t-1} cross-sectionally converts factors F_t into zero-investment long-short portfolios since $i_{1,t-1} + i_{2,t-1} + \cdots + i_{N,t-1} = 0$.

Appendix C. Firm Characteristics

This table lists the firm characteristics I include in the predictor matrix Z_t. Most are compiled by Green, Hand, and Zhang (2017).

Table C.1: Details of the Characteristics

No.	Firm characteristic	Author(s)	Year, Journal
1	size	Banz	1981, JFE
2	beta (weekly)	Fama & MacBeth	1973, JPE
3	beta (daily)	Fama & MacBeth	1973, JPE
4	idiosyncratic volatility (daily)	Ali, Hwang & Trombley	2003, JFE
5	beta squared (daily)	Fama & MacBeth	1973, JPE
6	beta squared (weekly)	Fama & MacBeth	1973, JPE
7	change in 6-month momentum	Gettleman & Marks	2006, WP
8	dollar trading volume	Chordia, Subrahmanyam & Anshuman	2001, JFE
9	idiosyncratic volatility (weekly)	Ali, Hwang & Trombley	2003, JFE
10	industry momentum	Moskowitz & Grinblatt	1999, JF
11	1-month momentum	Jegadeesh & Titman	1993, JF
12	6-month momentum	Jegadeesh & Titman	1993, JF
13	12-month momentum	Jegadeesh	1990, JF
14	36-month momentum	Jegadeesh & Titman	1993, JF
15	price delay	Hou & Moskowitz	2005, RFS

Table C.1: Details of the Characteristics (Continued)

No.	Firm characteristic	Author(s)	Year, Journal
16	share turnover	Datar, Naik & Radcliffe	1998, JFM
17	absolute accruals	Bandyopadhyay, Huang & Wirjanto	2010, WP
18	working capital accruals	Sloan	1996, TAR
19	# years since first Compustat coverage	Jiang, Lee & Zhang	2005, RAS
20	asset growth	Cooper, Gulen & Schill	2008, JF
21	cash flow to debt	Ou & Penman	1989, JAE
22	cash productivity	Chandrashekar & Rao	2009, WP
23	cash flow to price	Desai, Rajgopal & Venkatachalam	2004, TAR
24	cash flow to price (industry-adjusted)	Asness, Porter & Stevens	2000, WP
25	change in asset turnover (industry-adjusted)	Soliman	2008, TAR
26	change in shares outstanding	Pontiff & Woodgate	2008, JF
27	change in employees (industry-adjusted)	Asness, Porter & Stevens	1994, WP
28	change in inventory	Thomas & Zhang	2002, RAS
29	change in profit margin (industry-adjusted)	Soliman	2008, TAR
30	convertible debt indicator	Soliman	2008, TAR
31	current ratio	Ou & Penman	1989, JAE
32	depreciation PP&E	Holthausen & Larcker	1992, JAE
33	dividend initiation	Michaely, Thaler & Womack	1995, JF
34	dividend omission	Michaely, Thaler & Womack	1995, JF

Table C.1: Details of the Characteristics (Continued)

No.	Firm characteristic	Author(s)	Year, Journal
35	dividend to price	Litzenberger & Ramaswamy	1982, JF
36	growth in common shareholder equity	Richardson, Sloan, Soliman & Tuna	2005, JAE
37	earnings to price	Basu	1977, JF
38	gross profitability	Novy-Marx	2013, JFE
39	growth in capital expenditures	Anderson & Garcia-Feijoo	2006, JF
40	growth in long-term net operating assets	Fairfield, Whisenant & Yohn	2003, TAR
41	industry sales concentration	Hou & Robinson	2006, JF
42	employee growth rate	Bazdresch, Belo & Lin	2014, JPE
43	capital expenditures and inventory	Chen & Zhang	2010, JF
44	leverage	Bhandari	1988, JF
45	growth in long-term debt	Richardson, Sloan, Soliman & Tuna	2005, JAE
46	size (industry-adjusted)	Asness, Porter & Stevens	2000, WP
47	operating profitability	Fama & French	2005, JFE
48	organizational capital	Eisfeldt & Papanikolaou	2013, JF
49	% change in capital expenditures (industry-adjusted)	Abarbanell & Bushee	1998, TAR
50	% change in current ratio	Ou & Penman	1989, JAE
51	% in depreciation	Holthausen & Larcker	1992, JAE
52	% change in gross margin - % change in sales	Abarbanell & Bushee	1998, TAR
53	% change in quick ratio	Ou & Penman	1989, JAE

Table C.1: Details of the Characteristics (Continued)

No.	Firm characteristic	Author(s)	Year, Journal
54	% change in sales - % change in inventory	Abarbanell & Bushee	1998, TAR
55	% change in sales - % change in A/R	Abarbanell & Bushee	1998, TAR
56	% change in sales - % change in SG&A	Abarbanell & Bushee	1998, TAR
57	% change in sales-to-inventory	Ou & Penman	1989, JAE
58	percent accruals	Hafzalla, Lundholm & Van Winkle	2011, TAR
59	financial statements score	Piotroski	2000, JAR
60	quick ratio	Ou & Penman	1989, JAE
61	R&D increase	Eberhart, Maxwell & Siddique	2004, JF
62	R&D to market capitalization	Guo, Lev & Shi	2006, JBFA
63	R&D to sales	Guo, Lev & Shi	2006, JBFA
64	real estate holdings	Tuzel	2010, RFS
65	return on invested capital	Brown & Rowe	2007, WP
66	sales to cash	Ou & Penman	1989, JAE
67	sales to inventory	Ou & Penman	1989, JAE
68	sales to receivables	Ou & Penman	1989, JAE
69	secured debt	Valta	2016, JFQA
70	secured debt indicator	Valta	2016, JFQA
71	sales growth	Lakonishok, Shleifer & Vishny	1994, JF
72	sin stocks	Hong & Kacperczyk	2009, JFE

Table C.1: Details of the Characteristics (Continued)

No.	Firm characteristic	Author(s)	Year, Journal
73	sales to price	Barbee, Mukherji, & Raines	1996, FAJ
74	debt capacity/firm tangibility	Almeida & Campello	2007, RFS
75	tax income to book income	Lev & Nissim	2004, TAR
76	abnormal earnings announcement volume	Lerman, Livnat & Mendenhall	2007, WP
77	cash holdings	Palazzo	2012, JFE
78	change in tax expense	Thomas & Zhang	2011, JAR
79	corporate investment	Titman, Wei & Xie	2004, JFQA
80	earnings announcement return	Kishore, Brandt, Santa-Clara & Venkatachalam	2008, WP
81	number of earnings increase	Barth, Elliott & Finn	1999, JAR
82	return on assets	Balakrishnan, Bartov & Faurel	2010, JAE
83	earnings volatility	Francis, LaFond, Olsson & Schipper	2004, TAR
84	return on equity	Hou, Xue & Zhang	2015, RFS
85	revenue surprise	Kama	2009, JBFA
86	accrual volatility	Bandyopadhyay, Huang & Wirjanto	2010, WP
87	cash flow volatility	Huang	2009, JFE
88	financial statement score	Mohanram	2005, RAS
89	bid-ask spread	Amihud & Mendelson	1989, JF
90	illiquidity	Amihud	2002, JFM
91	maximum daily return	Bali, Cakici & Whitelaw	2011, JFE

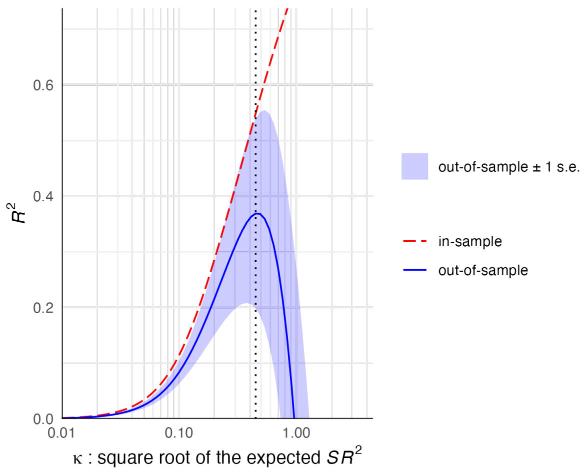
Table C.1: Details of the Characteristics (Continued)

No.	Firm characteristic	Author(s)	Year, Journal
92	return volatility	Ang, Hodrick, Xing & Zhang	2006, JF
93	volatility of liquidity (dollar trading volume)	Chordia, Subrahmanyam & Anshuman	2001, JFE
94	volatility of liquidity (share turnover)	Chordia, Subrahmanyam & Anshuman	2001, JFE
95	zero trading days	Liu	2006, JFE
96	book-to-market	Rosenberg, Reid & Lanstein	1985, JPM
97	book-to-market (industry-adjusted)	Asness, Porter & Stevens	2000, WP
98	skewness	Condard, Robert & Ghysels	2012, JF
99	kurtosis	Condard, Robert & Ghysels	2012, JF

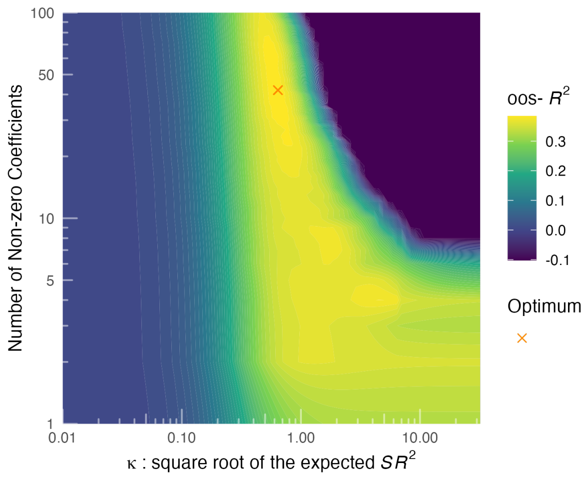
Appendix D. Figures

Figure D.1: Tradable Unevered Factors $F_{\hat{a}}$: OOS R^2 under Singular- and Dual-Penalty

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using 100 anomaly portfolios of daily tradable unlevered asset returns from 1970 to 2022. Panel (a) only employs L^2 penalty of which the strength is measured by prior root expected SR^2 (κ). Panel (b) also employs L^1 penalty of which the strength is measured by the number of retained factors. Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes of hyperparameters are ploted on logarithmic scale.



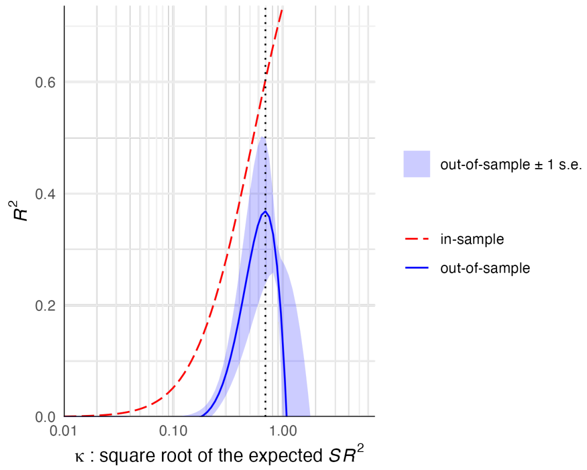
(a) L^2 penalty



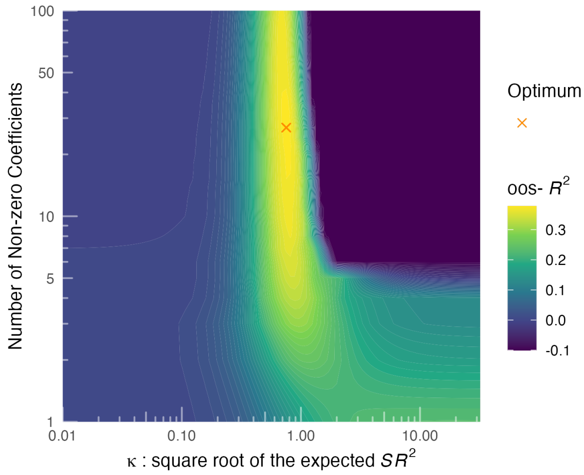
(b) L^1-L^2 penalties

Figure D.2: Principal Components P_e: OOS R^2 under Singular- and Dual-Penalty

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using the PCs of 100 anomaly portfolios of daily unlevered asset returns from 1951 to 2022. Panel (a) only employs L^2 penalty of which the strength is measured by prior root expected SR^2 (κ). Panel (b) also employs L^1 penalty of which the strength is measured by the number of retained factors. Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes of hyperparameters are ploted on logarithmic scale.



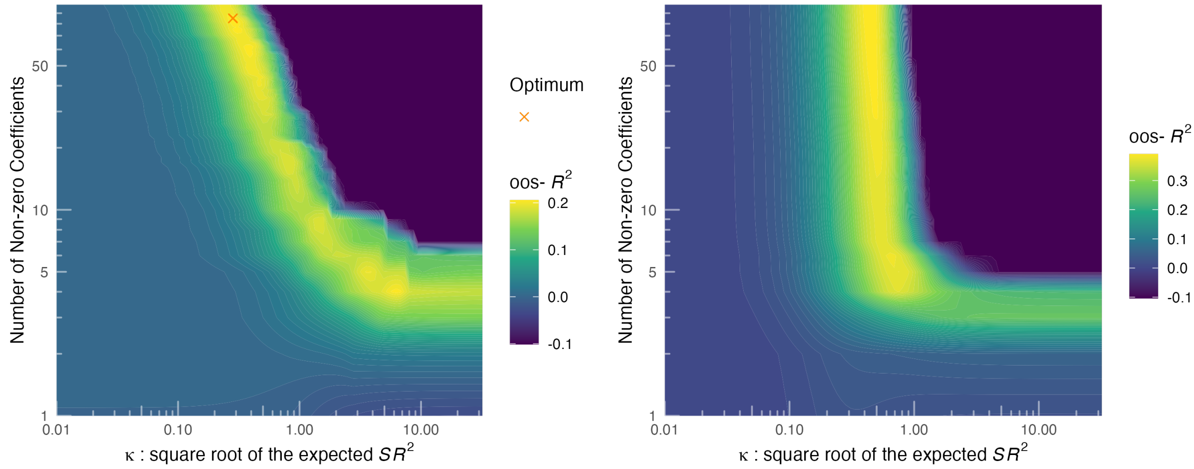
(a) L^2 penalty



(b) L^1-L^2 penalties

Figure D.3: OOS R^2 under Singular- and Dual-Penalty for Beta-Neutral Factors

This figure reports the OOS R^2 under different hyperparameters from 3-fold cross validation process using 100 anomaly portfolios. Panel (a) depicts the result for beta-neutral tradable levered factors ($\tilde{F}_{\hat{e}}$) and Panel (b) for beta-neutral PCs of levered factors ($\tilde{P}_e$). Hyperparameters corresponding to highest OOS R^2 are marked in the figure. Axes are ploted on logarithmic scale.



(a) tradable unlevered $\tilde{F}_{\hat{e}}$

(b) PCs of levered $\tilde{P}_e$