Frequency Dependent Risks in the Factor Zoo*

Jiantao Huang†

February 14, 2022

Abstract

I propose a novel framework to quantify frequency-dependent risks in the factor zoo. My approach generalizes canonical principal component analysis (PCA) by exploiting frequency-dependent information in asset returns. Empirically, the linear stochastic discount factor (SDF) composed of the first few low-frequency principal components (PCs) capture all the risk premium in asset returns. It also explains well the cross-section of characteristic-sorted portfolios. In contrast, high-frequency and canonical PCA have inferior performance since they fail to identify slow-moving information in asset returns. Moreover, I decompose the low-frequency SDF into two orthogonal priced components. The first component is constructed by high-frequency or traditional PCA. It is almost serially uncorrelated and relates to discount-rate news, intermediary factors, jump risk, and investor sentiment. The second component is slow-moving and captures business-cycle risks related to consumption and GDP growth. Hence, only low-frequency PCA identifies the second persistent component emphasized by many macro-finance models.

Keywords: Asset Pricing, Factor Models, Fourier transform, PCA.
JEL Classification Codes: C14, G11, G12.

*Any errors or omissions are my responsibility. I thank Thummim Cho, Ian Martin, Cameron Peng, Ran Shi, Dimitri Vayanos, and seminar participants at the London School of Economics, Tsinghua University PBC School, Chinese University of Hong Kong, University of British Columbia, Peking University HSBC Business School, and University of Hong Kong for helpful comments, discussions, and suggestions. I am particularly grateful to Christian Julliard, Dong Lou, and Svetlana Bryzgalova for their invaluable guidance and support.

†Department of Finance, London School of Economics, J.Huang27@lse.ac.uk

I Introduction

Explaining the cross-section of expected returns has been an important challenge in asset pricing literature. Researchers have acknowledged that the consumption-based capital asset pricing model (CCAPM)[1] provides little explanatory power, which has inspired a wide variety of new models. Some models introduce slow-moving components into the stochastic discount factor (SDF), such as the surplus consumption ratio in Campbell and Cochrane (1999) and the stochastic mean and variance of consumption growth in Bansal and Yaron (2004). In other models, the SDF consists only of fast-moving components, e.g., output jumps in Barro (2006), the intermediary's consumption growth in He and Krishnamurthy (2013), and sentiment-driven demand shocks in Kozak, Nagel, and Santosh (2018). Identifying the key determinants of the SDF, particularly the slow-moving components that are notoriously difficult to measure (see Alvarez and Jermann (2005)), remains an open question. This paper addresses this question through the lens of frequency-dependent risks. In addition, I seek to understand the frequency-specific drivers of expected returns and explore the role of distinct asset pricing models at different frequencies.

This paper generalizes canonical principal component analysis (PCA) to construct latent factors that explain the cross-section of monthly expected returns. The key novelty of my approach is that I exploit frequency-dependent information of asset returns to estimate latent factors. Using standard Fourier transform, I decompose the covariance matrix of monthly returns into high- and low-frequency components and estimate systematic factors in each frequency interval. I denote them as high- and low-frequency principal components (PCs) and use them as monthly tradable proxies for short- and long-term systematic risks.

When do frequency-dependent risks matter? I show that when asset returns are independent, high- and low-frequency latent factors are precisely identical to the canonical PCs. In other words, only when asset returns deviate from the independence assumption we need to study frequency-dependent risks. Empirically, low-frequency PCs contain a persistent element missed by high-frequency and conventional PCs. Moreover, this persistent missing part is essential in explaining expected returns and reflects business-cycle risks.

Asset pricing models often make parametric assumptions enforcing whether fast- or slow-moving economic shocks drive the SDF. Rather than assuming the existence of fast- or slow-moving elements, this paper lets the data speak and suggests that both two components are priced but reflect different economic fundamentals. The key for detecting the slow-moving component is the rich persistent information in the factor zoo. For example, Gupta and Kelly (2019) find that 48 of 65 investment anomalies have significantly positive AR(1)

[1] I refer to earlier versions of CCAPM developed by Rubinstein (1976), Lucas (1978), and Breeden (1979).

coefficients.[2] The low-frequency PCA boosts the signal of persistent information in the factor zoo and combines the factors' persistence into a few low-frequency PCs. Instead, the high-frequency or conventional PCA fails to detect them.

My empirical results are based on a large cross-section of 78 portfolios.[3] I divide the whole sample equally into two subsamples. I estimate the factor compositions and risk prices of frequency-specific PCs in the first subsample and examine their *out-of-sample (OOS)* performance in the second subsample. In the main analysis, the LF interval is between three and ten years, and I interpret it as the business-cycle frequency interval. In contrast, the HF interval is between zero and three years. The empirical findings are fourfold.

First, the SDF is sparse only in the space of low-frequency PCs. The low-frequency SDF comprising the seven largest low-frequency PCs is the "proper" benchmark: It yields an OOS Sharpe ratio of around 0.37 per month. Additional low-frequency PCs are redundant. In contrast, I need more than 20 high-frequency or canonical PCs to gain a comparable Sharpe ratio. Since high-frequency components account for 94% of time-series variations in asset returns, the large canonical PCs are virtually equivalent to the high-frequency latent factors. I also split the whole high-frequency interval into a few subintervals, but the SDF is dense even in the space of highly fast-moving factors (with a cycle length shorter than three months). Past research (e.g., Kozak, Nagel, and Santosh (2018, 2020)) often uses the first few PCs of single-period returns (identical to high-frequency PCs in the data) to construct the SDF. My paper shows that this standard practise can be improved by exploiting frequency-dependent information in asset returns.

Second, the low-frequency SDF cannot be explained by the high-frequency SDF or celebrated factor models in Fama and French (1993, 2015), Carhart (1997), and Hou, Xue, and Zhang (2015). Monthly alphas of the low-frequency SDF are significantly greater than 0.6%. In contrast, the low-frequency SDF can entirely span the high-frequency one. This evidence provides further justification for using the low-frequency SDF as the benchmark.

Third, I decompose the low-frequency SDF into fast- and slow-moving components. The first component is the optimal portfolio composed of high-frequency PCs. This SDF component is nearly identical to the SDF constructed by Kozak, Nagel, and Santosh (2018, 2020). I observe that the high-frequency SDF is almost serially uncorrelated and yields a monthly Sharpe ratio of 0.29, so I denote it as the fast-moving component. However, it still misses an essential slow-moving element. I project the low-frequency SDF into the space of the high-frequency SDF and extract an orthogonal part, denoted as the *missing-SDF*. This missing part, displaying a persistent dynamic according to the variance ratio test, explains 30% of

[2]The other 11 have positive yet insignificant coefficients. No factor has significantly negative coefficients.

[3]Test assets are long and short legs sorted by 39 firm features in Kozak, Nagel, and Santosh (2020).

the time-series variation of the low-frequency SDF and earns a monthly Sharpe ratio of 0.24.

Fourth, fast- and slow-moving components of the low-frequency SDF embody entirely different sources of economic risks. Precisely, the high-frequency SDF is correlated with market discount-rate news in Campbell and Vuolteenaho (2004), intermediary factors in He, Kelly, and Manela (2017), market jump risk proxied by the VXO index, and the sentiment-driven demand shocks from Baker and Wurgler (2006) investor sentiment. Instead, the slow-moving part of the SDF is related to consumption and GDP growth, which is great news for macro-finance. In fact, it also predicts the next quarter economic growth. Hence, the missing-SDF reflects slow-moving business-cycle risks.

My empirical findings have implications for asset pricing models, which link the SDFs to different economic fundamentals. Macro-finance models often use persistent shocks to macro variables, such as the stochastic mean of consumption growth, to magnify their prominence in the SDF. My paper confirms that asset returns carry useful persistent information related to macro fundamentals, but I can identify them only at low frequencies. My paper also reconciles the disconnection between asset returns and some macro fundamentals. For example, asset returns and consumption growth are almost uncorrelated at the quarterly frequency, so asset pricing seems to disconnect with the macroeconomy in short horizons. My paper confirms that the large PCs of short-horizon returns are unrelated to consumption growth. After removing high-frequency variations from asset returns, the remaining slow-moving component strongly correlates with macro fundamentals. Therefore, identifying the slow-moving component is salient for understanding and testing macro-finance models.

Furthermore, macro risks are not sufficient to explain the cross-section. The fast-moving component of the benchmark SDF commands a significant price of risk but is orthogonal to macro risks. Instead, the demand shocks from sentiment investors, the shocks to the intermediary sector, and market discount-rate news are essential in understanding the fast-moving component of the SDF (high-frequency SDF). Hence, different asset pricing models explain either fast- or slow-moving components of the SDF but not both.

There are two appealing benefits to studying asset returns at different frequencies. *First*, it helps to explore the dynamics of state variables in the SDF. I decompose the variance of an SDF (equivalently, the maximal achievable Sharpe ratio) into frequency-specific components. Also, I prove that if the SDF has a larger variance at high (low) frequencies, state variables entering the SDF are, on average, more fast-moving (slow-moving). Since a sparse low-frequency SDF embodies a significantly higher Sharpe ratio than a high-frequency one, slow-moving state variables are empirically more prominent than fast-moving ones.[4]

[4]The importance of a state variable X_t comes from the variance of X_t and its risk price squared (b_X^2). In latent factor models, I can identify only $\text{Var}(X_t)b_X^2$ rather than $\text{Var}(X_t)$ and b_X^2 individually.

Second, frequency-dependent PCA strengthens the signal of some systematic factors. Generally, a slow-moving (fast-moving) latent factor has a stronger signal at low (high) frequencies. Suppose a weak latent factor explains a tiny proportion of single-period returns.[5] In that case, the canonical PCA fails to identify it. However, frequency-specific PCA can recover this weak factor if its variance is large enough in a specific frequency interval. This paper shows that the low-frequency PCA recovers some essential priced weak factors with strong enough signals only at low frequencies. Instead, the high-frequency and canonical PCA identify them as idiosyncratic noises, so many small high-frequency and canonical PCs are needed to attain the same Sharpe ratio as a sparse low-frequency SDF.

Economic theory predicts the sparsity of latent factor models. The absence of near-arbitrage opportunities in Kozak, Nagel, and Santosh (2018) argues that only the largest PCs enter the SDF. However, this paper observes some small high-frequency (also canonical) PCs bringing nontrivial risk premia, so the absence of near-arbitrage opportunities fails. One explanation is that some economic shocks, such as the stochastic mean of consumption growth, are slow-moving and explain only a tiny fraction of single period returns. Hence, traditional PCA fails to detect these small but persistent shocks. Suppose market participants have Epstein-Zin preferences as in the long-run risk model. In that case, persistent shocks to economic fundamentals command sizable risk premia and constitute a considerable part of the SDF. Since the low-frequency PCA successfully captures these slow-moving elements, we observe the sparsity of the low-frequency SDF.

I.1 Closely Related Literature

This paper mainly contributes to two strands of literature. The first closely related branch of literature is the study of asset pricing models at different frequencies. We have known for a long time that both CAPM and CCAPM have better performance in the long horizon. For example, Handa, Kothari, and Wasley (1989) show that the size effect becomes statistically insignificant when the market beta is estimated using annual returns. Parker and Julliard (2005) measure ultimate consumption risk at a horizon of three years and document that it explains a large proportion of expected returns. Brennan and Zhang (2020) derive the CAPM with a stochastic investor horizon, and their estimates show that the probability distribution of investor horizons puts a massive weight on the interval between 8 and 20 months. Chernov, Lochstoer, and Lundeby (forthcoming) test asset pricing models using multi-horizon returns and report that single-period estimates of those models typically do a poor job of explaining long-term returns.

[5] Onatski (2012) and Lettau and Pelger (2020a) assume that the variance of a weak factor does not grow as the number of test assets converges to infinity.

However, all the above papers study factor models at a specific frequency instead of in a frequency interval. A few recent papers adopt spectral analysis to study frequency-dependent risks. First, Dew-Becker and Giglio (2016) study frequency-dependent risk prices in consumption-based models and show that only the long-run risk model can explain asset returns. Instead, my paper does not make a parametric assumption of the SDF. I construct the SDF using latent factors of asset returns and find that the SDF contains a huge fast-moving component that the consumption risk cannot explain.

Second, Bandi, Chaudhuri, Lo, and Tamoni (2021) use a Wold representation of the CAPM beta. Only the business cycle components within the frequency interval between 32 and 64 months can price the cross-sections. One key feature of their approach is assuming a vector autoregressive (VAR) process for state variables. In contrast, my paper takes a nonparametric point of view and is more robust to model misspecification of the state vector dynamics. In addition, we have different economic interpretations. Their paper claims that the business-cycle component of the market beta captures delayed price adjustments to new information in the market portfolio. Instead, my paper finds low-frequency systematic factors capture business-cycle risks, but short-term factors miss them.

Last but not least, Neuhierl and Varneskov (2021) decompose the covariance between asset returns and pricing factors via the Fourier transform and study the frequency-dependent risks. My paper improves their framework in a few aspects. Their paper studies factors individually, and their framework cannot handle the factor zoo. Instead, the framework in my paper is more suitable for the high-dimensional case. Also, they do not explore whether high- or low-frequency factors can explain the cross-section of average returns. Unlike their paper, I show that low-frequency latent factors are salient for cross-sectional asset pricing. The decomposition of the SDF into fast- and slow-moving components also improve our understanding of economic risks in the factor zoo.

The second branch of related literature is the abundant study of latent factor models after Ross (1976). Early empirical applications include Chamberlain and Rothschild (1983), Connor and Korajczyk (1986), Connor and Korajczyk (1988). Kozak, Nagel, and Santosh (2020) use PCA to estimate latent factors of a large cross-section of characteristic-managed portfolios and then estimate their risk prices via an elastic-net algorithm. Kelly, Pruitt, and Su (2019) propose the instrumented PCA to model both pricing errors and factor loadings as functions of firm characteristics, and they find that four IPCA factors explain the cross-section of individual stock returns. Lettau and Pelger (2020a) and Lettau and Pelger (2020b) generalize PCA by including a penalty term on the pricing errors in expected returns. Their method can identify weak factors with high Sharpe ratios, even when the canonical PCA omits them. This paper differs from previous literature in that I estimate latent factors

using frequency-dependent information in asset returns. As I show in Sections II and III, the importance of latent factors can change across frequencies, and the frequency-dependent PCA can also strengthen a factor's signal if it is not independent. Giglio and Xiu (2021) and Giglio, Xiu, and Zhang (2021) show that we can project a nontradable factor into the space of the largest several PCs of a huge cross-section of test assets. The risk premium of a nontradable factor is the expected return of its mimicking portfolio composed of the largest several PCs of single-period returns.

Nevertheless, I do not intend to develop a method that can outperform all previous forms of PCA. Instead, I aim to provide a novel framework that is suitable for analyzing frequency-dependent risks in the factor zoo. Moreover, my frequency-dependent PCA can also be integrated with other PCA methods. For example, we can construct the factor-mimicking portfolio composed of frequency-specific PCs and use the three-pass procedure in Giglio and Xiu (2021) to estimate the risk premium of nontradable factors.

II Methodology

Notation. $\mathbb{E}[\cdot]$, $\text{Var}[\cdot]$, and $\text{Cov}[\cdot]$ are the expectation, variance, and covariance operators. Suppose that $\boldsymbol{X_t}$ is an arbitrary $N \times 1$ vector of covariance-stationary random variables. $\boldsymbol{\mu_X}$, $\mathbb{E}_{t-1}[\boldsymbol{X_t}]$, and $\bar{\boldsymbol{X}}_{\boldsymbol{t}}$ denote the unconditional, conditional, and sample mean of $\boldsymbol{X_t}$. $\boldsymbol{\Sigma_X}(h)$ is the autocovariance matrix with lag h: $\boldsymbol{\Sigma_X}(h) = \mathbb{E}\big[\boldsymbol{X_{t+h}}\boldsymbol{X_t}\big] - \mathbb{E}[\boldsymbol{X_{t+h}}]\mathbb{E}[\boldsymbol{X_t}]^\top$. Particularly, $\boldsymbol{\Sigma_X}(0)$ is the unconditional covariance matrix, simply denoted by $\boldsymbol{\Sigma_X}$. $\hat{\boldsymbol{\Sigma}}_{\boldsymbol{X}}(h)$ is the sample estimate of $\boldsymbol{\Sigma_X}(h)$. $\text{Tr}\left[\boldsymbol{A}\right]$ is the trace of a matrix $\boldsymbol{A}$.

II.1 Asset Pricing Models

Suppose that there are N test assets, denoted by $\boldsymbol{R_t} = (R_{1t}, \ldots, R_{Nt})^\top$, and the sample size is T. This paper considers empirical applications in which both N and T are reasonably large, in particular, $\frac{N}{T} \to c < 1$. Motivated by the arbitrage pricing theory (APT) developed by Ross (1976), this paper studies an approximate factor pricing model, where the excess return on asset n, R_{nt}, is driven by a systematic component captured by K ($K < N$) latent factors and an idiosyncratic shock,

$$\underbrace{\boldsymbol{R_{t+1}}}_{N\times 1} = \underbrace{\boldsymbol{\alpha}}_{N\times 1} + \underbrace{\boldsymbol{\beta}}_{N\times K}\underbrace{\boldsymbol{F_{t+1}}}_{K\times 1} + \underbrace{\boldsymbol{e_{t+1}}}_{N\times 1}, \tag{1}$$

where $\boldsymbol{\alpha}$ denotes a vector of potential mispricings, $\boldsymbol{\beta F_{t+1}}$ is a vector of common components that are the product of factor loadings $\boldsymbol{\beta}$ and latent factors $\boldsymbol{F_{t+1}}$, and $\boldsymbol{e_{t+1}}$ is a vector of

idiosyncratic shocks. I further require $\boldsymbol{\beta F_{t+1}}$ and $\boldsymbol{e_{t+1}}$ to be orthogonal. Empirically, I need to estimate the common component and cannot identify $\boldsymbol{\beta}$ and $\boldsymbol{F_{t+1}}$ separately.

Moreover, I require only systematic risks, proxied by $\boldsymbol{F_{t+1}}$, to enter the SDF. In other words, this paper assumes a strong form of APT, whereas unsystematic risks $\boldsymbol{e_{t+1}}$ earn zero risk premia. Specifically, $\mathcal{M}_{t+1}$ is linear in factors $\boldsymbol{F_{t+1}}$,[6]

$$\mathcal{M}_{t+1} = 1 - \boldsymbol{b}^\top (\boldsymbol{F}_{t+1} - \boldsymbol{\mu_F}), \tag{2}$$

where $\boldsymbol{\mu_F}$ is the unconditional expectation of latent factors, and $\boldsymbol{b}$ is the vector of risk prices for systematic factors, capturing the compensation for bearing systematic risks. According to the Hansen and Jagannathan (1991) (**HJ**) bound, if $\boldsymbol{F}_{t+1}$ are tradable factors, $\boldsymbol{b}^\top \boldsymbol{F}_{t+1}$ is the mean-variance efficient (MVE) portfolio. Therefore, constructing the linear SDF is the equivalent of finding the MVE portfolio in the cross-section of test assets.

According to the fundamental asset pricing equation,

$$\mathbb{E}[\mathcal{M}_{t+1}\boldsymbol{R_{t+1}}] = \mathbb{E}\big\{\boldsymbol{R_{t+1}}[1 - \boldsymbol{b}^\top(\boldsymbol{F}_{t+1} - \boldsymbol{\mu_F})]\big\} = \boldsymbol{0_N} \tag{3}$$

$$\implies \mathbb{E}[\boldsymbol{R_{t+1}}] = \mathrm{Cov}(\boldsymbol{R_{t+1}}, \boldsymbol{F}_{t+1})\boldsymbol{b}, \tag{4}$$

so systematic risks, quantified by the covariance matrix, fully explain the cross-section of expected returns.

Past research documents the deviation of the independently and identically distributed (IID) assumption for asset returns. For instance, Chernov, Lochstoer, and Lundeby (forthcoming) calculate the variance ratio of the mean-variance efficient portfolios in notable factor models. According to their results, factor returns are far from IID. In addition, Haddad, Kozak, and Santosh (2020) show that the first few principal components of asset returns are predictable by their own portfolio-level log book-to-market ratio. Motivated by their findings, this paper deviates from the IID assumption of $\boldsymbol{R_{t+1}}$ by assuming that latent factors subsume all the time-series dependency. Specifically, I assume that a $p \times 1$ vector of mean-zero "latent" state variables $\boldsymbol{X_t}$ can predict factors $\boldsymbol{F_{t+1}}$ as follows:

$$\underbrace{\boldsymbol{F_{t+1}}}_{K\times 1} = \underbrace{\boldsymbol{\mu_F}}_{K\times 1} + \underbrace{\boldsymbol{\Phi_X}}_{K\times p}\underbrace{\boldsymbol{X_t}}_{p\times 1} + \underbrace{\boldsymbol{f_{t+1}}}_{K\times 1}, \tag{5}$$

where $\boldsymbol{\mu_F}$ is the unconditional mean of latent factors, $\boldsymbol{\Phi_X X_t}$ captures the time-varying conditional mean of latent factors, $\boldsymbol{f_{t+1}}$ is conditionally uncorrelated: $\mathbb{E}[\boldsymbol{f_{t+1}}] = \mathbb{E}_t[\boldsymbol{f_{t+1}}] =$

[6] I consider only excess returns in this paper, so the unconditional mean of $\mathcal{M}_{t+1}$ is unidentified. Without loss of generality, I normalize its mean to be one.

$\mathbf{0_K}$. Similarly, idiosyncratic shocks $\boldsymbol{e_{t+1}}$ are conditionally uncorrelated. Chamberlain and Rothschild (1983) also model idiosyncratic components as being cross-sectionally but not serially correlated. Since pricing errors are poorly predictable, such an assumption can be viewed as a good first-order approximation. I further plug equation (5) into the SDF,

$$\mathcal{M}_{t+1} = 1 - \boldsymbol{b}^\top \boldsymbol{f}_{t+1} - \boldsymbol{b}_{\boldsymbol{X}}^\top \boldsymbol{X_t}, \tag{6}$$

where $\boldsymbol{b}_{\boldsymbol{X}}^\top = \boldsymbol{b}^\top \boldsymbol{\Phi_X}$, and $\boldsymbol{b}_{\boldsymbol{X}}^\top \boldsymbol{X_t}$ drives the conditional mean of the SDF, capturing its full conditional dynamics. The formula for $\mathcal{M}_{t+1}$ in equation (6) relates to previous studies that decompose the SDF into permanent and transitory components (see Alvarez and Jermann (2005) and Hansen and Scheinkman (2009)).

In addition, Hansen, Heaton, and Li (2008) study parametric models of state variables, modelling them using a stationary vector autoregressive (VAR) model. Instead, this paper is agnostic about the state vector $\boldsymbol{X_t}$. $\boldsymbol{X_t}$ can be firm characteristics and macro indicators, such as book-to-market ratio and cay (see Lettau and Ludvigson (2001a)). By decomposing asset returns into frequency-dependent components, this paper can infer whether state variables critical in pricing the cross-section, on average, are more important at high or low frequencies. The next subsection introduces the Fourier transform as the non-parametric solution.

II.2 Frequency Domain Analysis

This paper uses the techniques in frequency domain analysis to model the time-series dependence of asset returns and decompose an empirical series into its repetitive or regular components. I start by motivating why the Fourier transform is a natural approach to study long-horizon asset returns. Suppose that an excess return process x_t follows an AR(1) process, $x_{t+1} = \rho_x x_t + \sqrt{1-\rho_x^2}\sigma_x \eta_{x,t+1}$, where ρ_x is the AR(1) coefficient, σ_x^2 is the unconditional variance, and $\eta_{x,t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$. When ρ_x is zero (negative, positive), the asset return follows an IID (fast-moving, slow-moving) process. When ρ_x is more positive, the asset return tends to be more persistent.

Figure A1 plots the cumulative returns in a 24-month rolling window for three AR(1) processes: $\rho_x \in \{-0.5, 0, 0.5\}$. No matter how persistent the time series is, its long-horizon return always exhibits a cyclical pattern. Hence, it is natural to project the long-horizon return on the sine and cosine functions: $x_{t,t+24} = a_0 + a_1 \sin(\frac{2\pi t}{48}) + a_2 \cos(\frac{2\pi t}{48}) + e_{t,t+24}$. Note that the deterministic processes $\sin(\frac{2\pi t}{48})$ and $\cos(\frac{2\pi t}{48})$ complete a cycle in 48 months, or equivalently, it has a cycle length of 48 months. Motivated by this observation, I can study the cyclical pattern of long-horizon asset returns by projecting them on the space of sine and cosine functions, and an M-month cumulative return corresponds to a cycle length of $2M$.

The frequency-domain analysis is the natural solution. Technically speaking, the Fourier transform decomposes a time series into orthogonal components at different frequencies. In the language of regression, it regresses the original time series into a sequence of sines and cosines functions.[7]

This paper uses ω to denote the frequency of a time-series process, which quantifies the number of cycles that this process completes per unit of time. Of equal interest is the period (or cycle length) of a time series, defined as the number of time points in a cycle: $\tau = \frac{1}{\omega}$. For instance, if ω is 0.1 in monthly data, the time series will finish 0.1 cycles in a month. Equivalently, it will take this process 10 months to complete one cycle.[8]

The spectral density matrix of $\boldsymbol{R_t}$ is defined as the Fourier transform of its auto-covariance matrices,

$$\boldsymbol{f_R}(\omega) = \sum_{h=-\infty}^{\infty} \boldsymbol{\Sigma_R}(h) \exp\{-2\pi i h\omega\}.$$

Through inverse Fourier transform, I can reverse engineer the auto-covariance matrix,

$$\boldsymbol{\Sigma_R}(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp\{2\pi i h\omega\} \boldsymbol{f_R}(\omega) d\omega.$$

In the study of asset pricing models, such as finding the tangency portfolio, investors focus on the covariance matrix of asset returns, that is, $h = 0$,

$$\text{Cov}(\boldsymbol{R_t}) := \boldsymbol{\Sigma_R} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathcal{R}(\boldsymbol{f_R}(\omega)) d\omega. \tag{7}$$

The Fourier transform of $\boldsymbol{R_t}$ decomposes asset returns as an equally weighted average of orthogonal components at different frequencies, so the covariance matrix of $\boldsymbol{R_t}$ equals the integral of the covariance matrix of its frequency-ω component, $\boldsymbol{f_R}(\omega)$, as in equation (7). In Appendix A1, I further show that only the real part of the spectral density matrix plays a role in estimating PCs. Hence, I focus on $\mathcal{R}(\boldsymbol{f_R}(\omega))$, the real part of the spectral density. Equation (7) also implies that $\mathcal{R}(\boldsymbol{f_R}(\omega))$ is the contribution to the covariance matrix from the frequency-ω component. If test asset returns are IID, the spectral density matrix of asset returns is constant across frequencies; that is $\boldsymbol{f_R}(\omega) = \boldsymbol{\Sigma_R}$ for every ω.

Why should we study the frequency-specific covariance matrices of asset returns? One reason is that the single-period covariance matrix often fails to capture systematic risks

[7] According to the spectral representation theorem in Hannan (2009), a covariance-stationary time series can be approximated by a sum of sine and cosines random variables with different variances across frequencies. (see appendix A.1).

[8] In addition, the absolute value of ω is no larger than 0.5 since any time series spends at least two months completing a cycle.

critical in explaining risk premia. For example, Brennan and Zhang (2020) show that yearly CAPM beta, which equals the covariance between annualized asset returns and the market portfolio, can explain the cross-section of 25 Fama-French size-B/M monthly portfolios. In contrast, the monthly CAPM beta entirely fails. Therefore, the single-period covariance, like the single-period CAPM beta, possibly misspecifies actual systematic risks, rendering the estimation of risk prices difficult or even impossible. This observation calls for the study of frequency-dependent systematic risks.

I estimate the spectral density matrix via discrete Fourier transform (DFT).[9] A simple example is in Figure A3, where a deterministic time series x_t in panel (c) consists of two components. The first component in panel (a) is slow-moving, with a frequency equal to 0.05, which completes a cycle every 20 periods. Another component of x_t in panel (b) is fast-moving, spending only two periods repeating a cycle. As in panel (d), DFT decomposes the variance of x_t into two parts contributed by low-frequency and high-frequency fluctuations.

Like other non-parametric estimation methods, the DFT estimate of the spectral density matrix at a particular frequency is susceptible to significant uncertainties. To reduce the variance, I divide the frequency intervals into three groups and estimate the spectral density matrix in each frequency interval.

What are the ideal cutoff points of the entire frequency interval? Past research can give some hints on this question. Dew-Becker and Giglio (2016) derive a closed-form solution to frequency-specific risk prices of parametric CCAPMs. They observe that only the long-run risk with cycles more prolonged than the business cycle is priced in the cross-section. In addition, Bandi, Chaudhuri, Lo, and Tamoni (2021) use the orthogonal Wold decomposition of CAPM's beta, and they find that only the business-cycle component of CAPM beta in the frequency interval between 32 and 64 months is priced. These observations motivate the following division of frequency intervals.

More specifically, I consider the following divisions of frequency intervals: (1) $\tau = \frac{1}{\omega} < 36$ months (high-frequency, denoted as HF), (2) $\tau = \frac{1}{\omega} \in [36, 120]$ months (low-frequency, or business cycle frequency, denoted as LF), and (3) $\tau = \frac{1}{\omega} > 120$ months (Above-LF, or A-LF). This paper considers the second frequency interval as being closely related to business cycles. Generally, the covariance matrix in the third group, with a cycle length greater than 120 months, is difficult to estimate non-parametrically. In later analysis, I also divide the HF interval into several sub-intervals. In addition, I consider alternative LF intervals in the

[9]Details about DFT can be found in Appendix A.2.

robustness check. With the above division, I decompose the covariance matrix of $\boldsymbol{R_t}$,

$$\boldsymbol{\Sigma_R} = \int_{\omega\in\Omega_{HF}} \mathcal{R}(\boldsymbol{f_R}(\omega))d\omega + \int_{\omega\in\Omega_{LF}} \mathcal{R}(\boldsymbol{f_R}(\omega))d\omega + \int_{\omega\in\Omega_{A-LF}} \mathcal{R}(\boldsymbol{f_R}(\omega))d\omega \tag{8}$$

$$= |\Omega_{HF}|\boldsymbol{\Sigma}_{\boldsymbol{R}}^{HF} + |\Omega_{LF}|\boldsymbol{\Sigma}_{\boldsymbol{R}}^{LF} + |\Omega_{A-LF}|\boldsymbol{\Sigma}_{\boldsymbol{R}}^{A-LF}, \tag{9}$$

where Ω_{HF}, Ω_{LF}, and Ω_{A-LF} denote the set of HF, LF, and Above-LF, with lengths $|\Omega_{HF}|$, $|\Omega_{LF}|$ and $|\Omega_{A-LF}|$ ($|\Omega_{HF}| + |\Omega_{LF}| + |\Omega_{A-LF}| = 1$).

Proposition 1 (Decomposition of asset returns' spectral density matrix) *I assume that $\boldsymbol{e_{t+1}}$ and $\boldsymbol{f_{t+1}}$ are conditional uncorrelated, and they are orthogonal. Then the spectral density matrices of $\boldsymbol{e_{t+1}}$ and $\boldsymbol{f_{t+1}}$ are constant across frequencies and equal to their unconditional covariance matrices $\boldsymbol{\Sigma_e}$ and $\boldsymbol{\Sigma_f}$, respectively. Moreover, I can decompose the spectral density matrix of $\boldsymbol{R_t}$ as,*

$$\boldsymbol{f_R}(\omega) = \boldsymbol{\beta}\boldsymbol{f_F}(\omega)\boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e} = \boldsymbol{\beta}\boldsymbol{\Sigma_f}\boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e} + \boldsymbol{\beta_X}\boldsymbol{f_X}(\omega)\boldsymbol{\beta_X}^\top, \tag{10}$$

where $\boldsymbol{f_F}(\omega)$ and $\boldsymbol{f_X}(\omega)$ are the spectral density matrices of latent systematic factors and state variables, and $\boldsymbol{\beta_X} = \boldsymbol{\beta}\boldsymbol{\Phi_X}$.

A simple derivation of proposition 1 is in Appendix B.1. A key observation in equation (10) is that only the last component related to state variables is frequency-dependent, as $\boldsymbol{e_{t+1}}$ and $\boldsymbol{f_{t+1}}$ are conditionally uncorrelated. Furthermore, if I estimate latent factors of asset returns at different frequencies, I can study the difference between HF and LF systematic risks. More precisely, equation (10) indicates that the dynamics of state variables entirely drive the difference between HF and LF systematic risks. Similarly, I also decompose the SDF into frequency-dependent components and illustrate how the maximal Sharpe ratio implied by the SDF varies across frequencies.

Proposition 2 (Spectral density function of the SDF) *I normalize latent state variables $\boldsymbol{X_t}$ such that they are uncorrelated. Define risk prices of $\boldsymbol{X_t}$ as $\boldsymbol{b_X} = \boldsymbol{\Phi_X}^\top\boldsymbol{b}$: $\boldsymbol{b_X} = (b_{X,1}, \ldots, b_{X,p})^\top$. Then the unconditional variance of the SDF is*

$$\mathrm{Var}(\mathcal{M}_{t+1}) = \boldsymbol{b}^\top\boldsymbol{\Sigma_f}\boldsymbol{b} + \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j=1}^{p} b_{X,i}^2 f_{X_i}(\omega)d\omega, \tag{11}$$

and the spectral density function of $\mathcal{M}_{t+1}$ is

$$f_{\mathcal{M}}(\omega) = \boldsymbol{b}^\top\boldsymbol{\Sigma_f}\boldsymbol{b} + \sum_{j=1}^{p} b_{X,i}^2 f_{X_i}(\omega), \tag{12}$$

where $f_{X_i}(\omega)$ is the spectral density function of the i-th state variable X_i.

I derive proposition 2 in Appendix B.2. Since state variables are latent, I can always normalize them such that they are uncorrelated. An alternative interpretation of the normalization of $\boldsymbol{X_t}$ is that latent state variables are PCs of conditional expectations of factors. The above derivation shows that the maximal Sharpe ratio of the economy is frequency dependent. Moreover, the spectral density function of $\mathcal{M}_{t+1}$, denoted by $f_{\mathcal{M}}(\omega)$, varies across frequencies only due to the second term $\sum_{j=1}^{p} b_{X,i}^2 \boldsymbol{f_{X_i}}(\omega)$. I interpret this quantity as the weighted-average spectral density function of latent state variables, with weights proportional to the squared risk prices of state variables. If, on average, $\sum_{j=1}^{p} b_{X,i}^2 \boldsymbol{f_{X_i}}(\omega)$ is larger at high (low) frequencies, it implies that high (low) frequency information is more prominent in this cross-section of asset returns.[10] In addition, the spectral decomposition of $\mathcal{M}_{t+1}$ can only identify the state variable with a non-zero price of risk $b_{X,i}$.

Similar to this paper, Neuhierl and Varneskov (2021) use the Fourier transform to study the dynamics of state variables driving asset returns. They also show how to map their SDF into some canonical consumption-based asset pricing models. While the spectral density function of the SDF in IID CCAPM is constant across frequencies, other candidate models such as the long-run risk model in Bansal and Yaron (2004) have persistent SDFs. From the theoretical point of view, the LF component is more critical than the HF one in the SDF. A limitation of Neuhierl and Varneskov (2021) is that they consider only a single factor and explore how its risk premium varies across frequency. Differently, this paper aims to handle a factor zoo and extract frequency-dependent systematic risks in the large cross-section using the techniques introduced in the following subsections.

Remark 3 (Interpretation of the Frequency) *The frequency is different from the turnover of a factor strategy. Let us consider two value strategies: the first is the monthly rebalanced HML, or HML devil, from the AQR library. The second strategy is the yearly rebalanced HML from Ken French library. Even though these two strategies have different turnovers, their correlation is as high as 0.9. Furthermore, I compute their autocorrelations and variance ratios — these two HML strategies have pretty similar patterns.*

[10] The prominence of a state variable X_t comes from two sources: the variance of X_t and its risk price squared (b_X^2). Since I use PCs as factors and state variables are latent, I can identify only $\text{Var}(X_t)b_X^2$ rather than $\text{Var}(X_t)$ and b_X^2 individually.

II.3 Estimation of Risk Factors

Under the assumption that $\boldsymbol{F_{t+1}}$ and $\boldsymbol{e_{t+1}}$ are orthogonal, I can represent the covariance matrix of asset returns as following:

$$\boldsymbol{\Sigma_R} = \boldsymbol{\beta}\ \boldsymbol{\Sigma_F}\ \boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e}. \tag{13}$$

Systematic factors $\boldsymbol{F_{t+1}}$ are not directly observable, so I aim to estimate tradable proxies for them. A common way to estimate model (1) is the Principal Component Analysis (PCA), which relies on the eigendecomposition of $\boldsymbol{\Sigma_R}$,

$$\boldsymbol{\Sigma_R} = \boldsymbol{Q\Lambda Q}^\top, \text{ with } \boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \ldots, \lambda_N\}, \tag{14}$$

where $\boldsymbol{Q}$ is the matrix of eigenvectors ($\boldsymbol{Q}^\top\boldsymbol{Q} = \boldsymbol{I_N}$), and $\boldsymbol{\Lambda}$ is the diagonal matrix of eigenvalues in a descending order. A common practice of PCA is to estimate $\boldsymbol{\beta}$ as the first K columns of $\boldsymbol{Q}$, denoted by $\boldsymbol{Q_K}$. Moreover, the estimates of K principal components (PC) are $\hat{\boldsymbol{F}}_{\boldsymbol{t}} = \boldsymbol{Q}_{\boldsymbol{K}}^\top \boldsymbol{R_t}$, uncorrelated by construction.

This paper uses the normalization $\boldsymbol{\beta}^\top\boldsymbol{\beta} = \boldsymbol{I_K}$ and $\boldsymbol{\Sigma_F} = \text{diag}\{\sigma^2_{F,1}, \ldots, \sigma^2_{F,K}\}$ in all following analyses, with exceptions unless stated. Furthermore, I allow for a mixture of strong and weak factors. In fact, most of asset pricing studies, as in this paper, use diversified portfolios as test assets, and the number of test assets is not truly infinite. In this paper, I differentiate strong and weak factors based on their variances.

Bai (2003) proves the asymptotic consistency of PCA when all factors are strong factors, which affect an increasing number of test asset returns as N goes to infinity. Mathematically, $\frac{\sum_{t=1}^{T} \boldsymbol{F_t}\boldsymbol{F_t}^\top}{T} \to \boldsymbol{\Sigma_F}$ and $\frac{\boldsymbol{\beta}^\top\boldsymbol{\beta}}{N} \to \boldsymbol{\Sigma_N}$, where both $\boldsymbol{\Sigma_F}$ and $\boldsymbol{\Sigma_N}$ are positive definite matrices. Bai and Ng (2002) make the same assumption and propose an asymptotically consistent algorithm to determine the number of latent factors in model (1). Under the normalization chosen by this paper, the above assumption is the equivalent of explosive eigenvalues of $\boldsymbol{\Sigma_F}$. That is to say, the largest K eigenvalues of $\boldsymbol{\Sigma_R}$ and $\boldsymbol{\Sigma_F}$ will go to infinity as $N \to \infty$.

In addition, Bai (2003), like other papers, allows the idiosyncratic terms $\boldsymbol{e_t}$ to be only weakly correlated both cross-sectionally and over time. Furthermore, idiosyncratic shocks explain a finite amount of time-series variations in asset returns; that is

$$\lim_{N,T\to\infty} \sup \text{ max eval}\big(\boldsymbol{\Sigma_e}\big) < \infty,$$

where max eval($\boldsymbol{A}$) denotes the maximal eigenvalue of matrix $\boldsymbol{A}$.

In empirical applications, however, it is uncommon to come across the ideal case in which we can clearly separate large eigenvalues related to latent factors from small eigenvalues rep-

resenting idiosyncratic shocks. A few papers have documented that PCA cannot consistently estimate model (1) when some latent factors are weak (see Onatski (2012), Lettau and Pelger (2020a)). Contrary to a strong factor, a weak one explains a relatively smaller fraction of time-series variations in asset returns. Alternatively, we can interpret a weak factor as having a finite variance or a relatively small variance in a finite sample.

Some weak factors are necessary to explain the cross-section of asset returns. For example, Lettau and Pelger (2020b) show that the omission of weak factors with a high Sharpe ratio can deteriorate the performance of latent factor models. However, the question is, when will PCA ignore the weak factors?

Benaych-Georges and Nadakuditi (2011) shed light on this question. Suppose that the covariance matrix of asset returns can be decomposed into the sum of two matrices as in equation (13), and one of them, such as $\boldsymbol{\Sigma_e}$, has bounded eigenvalues. Under this setup, the k-th ($k \leq K$) eigenvalue of $\boldsymbol{\Sigma_R}$, representing the k-th systematic factor F_{kt}, is identified if the k-th eigenvalue of $\boldsymbol{\beta}\ \boldsymbol{\Sigma_F}\ \boldsymbol{\beta}^\top$, equal to $\sigma^2_{F,k}$ under the normalization, is greater than a certain threshold; that is

$$\lambda_k\big(\boldsymbol{\beta}\ \boldsymbol{\Sigma_F}\ \boldsymbol{\beta}^\top\big) = \sigma^2_{F,k} > \lambda_{crit},$$

where $\lambda_k(\boldsymbol{A})$ denotes the k-th largest eigenvalue of matrix $\boldsymbol{A}$, and λ_{crit} is related to the limit of $\frac{N}{T}$ and the upper bound of eigenvalues for $\boldsymbol{\Sigma_e}$. Otherwise, a phenomenon called eigenvalue phase transition occurs, and the factor k is no longer identified. Now let us look at a simple example.

Example 4 (Single-factor model) *Suppose that there is only one systematic factor in model (1), and the idiosyncratic vector* $\boldsymbol{e_t}$ *has a covariance matrix* $\sigma^2\boldsymbol{I_N}$ *($\sigma^2 < \infty$). I further normalize the factor loadings such that* $\boldsymbol{\beta}^\top\boldsymbol{\beta} = 1$*, and the variance of* $\boldsymbol{F_t}$ *is* σ^2_F *($\sigma^2_F < \infty$). As* $\frac{N}{T} \to c < 1$*, the distribution of eigenvalues for* $\mathrm{Var}(\boldsymbol{e_t})$ *converges to the Marchenko–Pastur distribution, with lower and upper bounds* $\sigma^2(1-\sqrt{c})^2$ *and* $\sigma^2(1+\sqrt{c})^2$*.*

According to corollary 2 in Lettau and Pelger (2020a), when $\sigma^2_F \leq \sqrt{c}\sigma^2$*, the top eigenvalue converges to* $\sigma^2(1+\sqrt{c})^2$*. Consequently, PCA can no longer identify* $\boldsymbol{F_t}$*, and the correlation between true factor* $\boldsymbol{F_t}$ *and the PCA estimate* $\hat{\boldsymbol{F}}_{\boldsymbol{t}}$ *converges to zero.*

A strong factor has a variance that is much more considerable than the critical point at all frequencies, so it is always identifiable. However, there are some "marginal factors" whose signals are strong enough only at high or low frequencies. It depends on the dynamics of state variables driving this factor. Suppose that a weak factor in example 4 has a variance slightly less than the critical value $\sqrt{c}\sigma^2$, but it follows an AR(1) process: $F_t = \rho_F F_{t-1} + \sqrt{1-\rho_F^2}e_{F,t},\ e_{F,t} \overset{\text{iid}}{\sim} \mathcal{N}(0,\sigma_F^2)$. If ρ_F is more positive (negative), F_t is more slow-moving (fast-moving). The spectral density of F_t is in Figure A2. For instance, when ρ_F is 0.5, the variance

of F_t at low frequencies is roughly four times the unconditional variance. It is possible that a weak factor is unidentified by canonical PCA but stands out at low frequencies if its signal is persistent and strong enough in the long horizon. This observation also motivates the frequency-dependent PCA.

Definition 1 (Frequency-dependent PCA) *Suppose that* $\boldsymbol{\Sigma_R^{HF}}$, $\boldsymbol{\Sigma_R^{LF}}$, *and* $\boldsymbol{\Sigma_R^{A-LF}}$ *are high-, low-, and above-low-frequency covariance matrices of the* N*-dimensional random vector* $\boldsymbol{R_t}$. *The eigendecomposition of* $\boldsymbol{\Sigma_R^Z}$, $Z \in \{HF, LF, A\text{-}LF\}$, *is*

$$\boldsymbol{\Sigma_R^Z} = \boldsymbol{Q}^Z \boldsymbol{\Lambda}^Z (\boldsymbol{Q}^Z)^\top, \text{ with } \boldsymbol{\Lambda}^Z = diag\{\lambda_1^Z, \ldots, \lambda_N^Z\},$$

where $\boldsymbol{Q}^Z$ *are eigenvectors of* $\boldsymbol{\Sigma_R^Z}$, *that is,* $(\boldsymbol{Q}^Z)^\top \boldsymbol{Q}^Z = \boldsymbol{I_N}$, *and* $\boldsymbol{\Lambda}^Z$ *is the diagonal matrix of eigenvalues in descending order. Define the latent factors in the frequency* $Z \in \{HF, LF, A\text{-}LF\}$ *as* $\boldsymbol{F_t^Z} = (\boldsymbol{Q}^Z)^\top \boldsymbol{R_t}$.

Intuitively, I rotate the space of canonical PCs to target the short-term and long-term common variations in asset returns. In other words, frequency-dependent PCA aims to select monthly proxies for short-term and long-term systematic risks. A special case is when asset returns are IID. Since $\boldsymbol{\Sigma_R^Z}$ are identical in this case, PCA, HF-PCA, and LF-PCA will deliver the exact estimates across frequencies.

II.4 Estimation of Risk Prices

This paper always uses principal components of asset returns as systematic factors, $\boldsymbol{F_t} = (\boldsymbol{Q}^Z)^\top \boldsymbol{R_t}$, $Z \in \{\text{HF}, \text{LF}, \text{A-LF}\}$. Since $\boldsymbol{F_t}$ are always tradable, estimating the linear SDF in equation (2) is the same as finding the optimal portfolio weights, $\boldsymbol{b}$, such that $\boldsymbol{b}^\top \boldsymbol{F_t}$ is the MVE portfolio. If the SDF prices the cross-section of asset returns, it also prices all tradable factors, so I can rewrite equation (4) as follows:

$$\boldsymbol{\mu_F} = \boldsymbol{\Sigma_F} \boldsymbol{b}. \tag{15}$$

I can solve the risk prices from equation (15), $\boldsymbol{b} = \boldsymbol{\Sigma_F^{-1}} \boldsymbol{\mu_F}$. Therefore, risk prices $\boldsymbol{b}$ are proportional to the MVE portfolio weights. In a finite sample, I need to estimate both $\boldsymbol{\Sigma_F}$ and $\boldsymbol{\mu_F}$. Past research have shown that a naive estimator such as $\hat{\boldsymbol{b}} = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{F}}^{-1} \bar{\boldsymbol{F}}_{\boldsymbol{t}}$ does not perform well in real datasets. For example, Tu and Zhou (2011) show that the estimated Markowitz (1952) portfolio not only underperforms the naive 1/N rule, in which investors invest equally across N assets, but also earns negative risk-adjusted returns. Kozak, Nagel, and Santosh (2020) argue that the majority of uncertainty comes from the estimation of factor means $\boldsymbol{\mu_F}$ and propose a simple Bayesian procedure to estimate $\boldsymbol{b}$.

To compare with Kozak, Nagel, and Santosh (2020), this paper adopts a similar strategy, which assumes that the covariance matrix of factor returns, $\boldsymbol{\Sigma_F}$, is known and focuses on the modelling of mean factor returns. Furthermore, equation (15) does not hold exactly in finite sample, so I include pricing errors $\boldsymbol{\alpha}$ on the right-hand side of equation (15): $\boldsymbol{\mu_F} = \boldsymbol{\Sigma_F b} + \boldsymbol{\alpha}, \quad \boldsymbol{\alpha} \sim \mathcal{N}(\mathbf{0_N}, \frac{1}{T}\boldsymbol{\Sigma_F})$. Finally, I assign a normal prior for risk prices: $\boldsymbol{b} \sim \mathcal{N}(\mathbf{0_K}, \frac{\kappa^2}{\tau}\boldsymbol{I_K}),\ \tau = \text{Tr}\left[\boldsymbol{\Sigma_F}\right]$. Under such a prior distribution, the prior expectation on the squared Sharpe ratio of factor returns is equal to

$$\mathbb{E}_{prior}[SR_F^2] = \mathbb{E}_{prior}[\boldsymbol{b}^\top \boldsymbol{\Sigma_F b}] = \frac{\kappa^2}{\tau}\,\text{Tr}\left[\boldsymbol{\Sigma_F}\right] = \kappa^2.$$

Also, the posterior distribution of $\boldsymbol{b}$, conditional on $(\boldsymbol{\mu_F}, \boldsymbol{\Sigma_F})$, is

$$p(\boldsymbol{b} \mid \boldsymbol{\mu_F}, \boldsymbol{\Sigma_F}) \propto \exp\left\{-\frac{T}{2}\left[(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b})^\top \boldsymbol{\Sigma_F}^{-1}(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b}) + \frac{\tau}{\kappa^2 T}\boldsymbol{b}^\top \boldsymbol{b}\right]\right\}.$$

Therefore, the posterior mode of $\boldsymbol{b}$ is the solution to the below objective function I,

$$\min_{\boldsymbol{b}}\left\{(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b})^\top \boldsymbol{\Sigma_F}^{-1}(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b}) + v_2 \boldsymbol{b}^\top \boldsymbol{b}\right\}, \tag{16}$$

where $v_2 = \frac{\text{Tr}\left[\boldsymbol{\Sigma_F}\right]}{\mathbb{E}_{prior}[SR_F^2]\times T}$. A detailed derivation of equation (16) is in Appendix B.3. For simplicity, I will denote $\sqrt{\mathbb{E}_{prior}[SR_F^2]}$ as SR_{prior}, or simply call it the prior Sharpe ratio. However, due to Jensen's inequality, $\mathbb{E}_{prior}[SR_F] \leq \sqrt{\mathbb{E}_{prior}[SR_F^2]}$, so SR_{prior} is an upper bound on the expected Sharpe ratio under prior distribution. Objective function I is to minimize squared Sharpe ratio of pricing errors (or equivalently maximize R^2_{GLS}), subject to L_2-penalty. In addition, I include factors into the model based on their ability of explaining time-series variations. That is to say, I include the first K largest PCs into analysis when I consider a $K-$factor model.

Kozak, Nagel, and Santosh (2020) extends equation (16) by including the L_1-penalty,

$$\min_{\boldsymbol{b}}\left\{(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b})^\top \boldsymbol{\Sigma_F}^{-1}(\boldsymbol{\mu_F} - \boldsymbol{\Sigma_F b}) + 2v_1|\boldsymbol{b}|_1 + v_2 \boldsymbol{b}^\top \boldsymbol{b}\right\}. \tag{17}$$

Since the principal components are uncorrelated by constructions, its covariance matrix $\boldsymbol{\Sigma_F}$ is diagonal with elements equal to eigenvalues of the covariance matrix of test assets. The closed-form solution to optimization problem (17) is

$$\hat{\lambda}_{i,KNS} = \begin{cases} \frac{\mu_{F,i} - v_1}{\sigma^2_{F,i} + v_2}, & \text{if } \mu_{F,i} \geq v_1, \\ 0, & \text{if } \mu_{F,i} < v_1, \end{cases} \tag{18}$$

so the above algorithm selects a certain factor j whenever it has a mean greater than v_1. In other words, v_1 controls the sparsity of factor models. Moreover, factors with small variances are shrunk more heavily by the L_2-penalty. This makes sense as those factors are more likely to be idiosyncratic risks that should not command sizeable risk premia.

This paper shows the empirical results using both objective functions (16) and (17). Suppose systematic factors that explain a large amount of time-series variations can capture most of the risk premium. In that case, estimates by (16) or (17) should be largely similar. It also implies that we can find a sparse factor model consisting of large PCs of asset returns. Finally, I emphasize that I also assume there is only one true SDF. One of this paper's main objectives is to determine whether the SDF comprised of canonical, HF, or LF PCs is a better approximation to the tangency portfolio.

III Empirics

I now proceed to the empirical studies. The first step is to estimate frequency-specific risk factors using the techniques introduced in Section II and to investigate whether the SDF composed of HF or LF factors is a better approximation to the tangency portfolio. Next, I explore whether some celebrated factor models proposed in the literature, such as Fama-French three factors, can explain these SDFs. Finally, I attempt to understand the economic fundamentals behinds SDFs. I begin this section with the analysis of 25 Fama-French portfolios to show how the factor structure of asset returns varies across frequencies. I then carry out the main analysis in a large cross-section of portfolios, studying which frequency is salient for the cross-sectional asset pricing.

III.1 Sample and Data

My primary data source comes from the characteristic-managed portfolios in Kozak, Nagel, and Santosh (2020). The definition of firm characteristics and the data are on Serhiy Kozak's website. There are 51 firm characteristics in Kozak, Nagel, and Santosh (2020),[11] but I select 39 of them to ensure the sample size large enough to estimate the low-frequency covariance matrix of asset returns. In the benchmark analysis, I split the sample from August 1963 to December 2019 into two halves and focus on the out-of-sample performance, which imposes additional challenges on the estimation due to the smaller subsamples. Firm characteristics can be further categorized into eight groups, as in Table A2. Kozak, Nagel, and Santosh (2020) also exclude stocks with market equity below 0.01% of the aggregate US market

[11] I thank the authors for sharing the data on their website. A more specific description of how to construct this dataset can be found on Serhiy Kozak's website: https://sites.google.com/site/serhiykozak/data.

cap, alleviating the impact of microcaps. Each month, individual stocks are sorted into 10 portfolios based on each of the 39 firm characteristics. They construct portfolios with weights equal to cross-sectional ranks of a given stock's characteristic, which is centered and normalized by the sum of absolute values of all ranks in the cross-section.

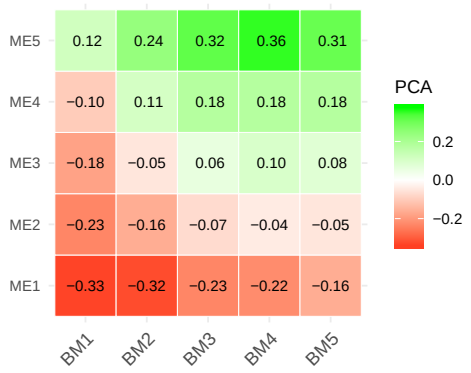
I also use the Fama-French 25 size and book-to-market (total variance) portfolios in Section III.2. I download the data from Ken French's library. In succeeding analysis, I supplement the main dataset with additional economic variables. Detailed variable definitions and data sources are provided in Table A1.

III.2 Starting Examples: 25 Fama-French Portfolios

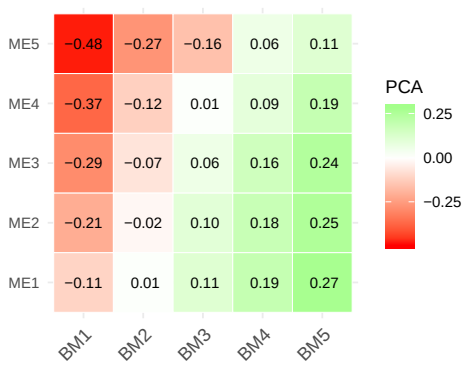
To illustrate how the factor structures vary across frequencies, I start with the 25 Fama-French size and book-to-market portfolios. The numbers in Figure 1 are 25 portfolios' factor loadings, equivalently their portfolio weights. The sample spans from August 1963 to December 2019. In each graph, the x-axis shows five buckets of book-to-market ratio, whereas the y-axis plots five levels of firm size. For instance, ME1 represents small firms, and BM5 means high book-to-market portfolios. Since PC1 is always the level (identically market) factor, I will display only the second and third PCs. In addition, I note that PCs and HF-PCs are almost identical; therefore, I will focus on explaining the difference between HF-PCs and LF-PCs.

First of all, let us look at Panel (c): In each column, the second HF-PC positively loads on all large portfolios in ME5 but negatively on small portfolios in ME1. Therefore, HF-PC2 is a size factor. Next, in Panel (d), the portfolio weights of all five portfolios in BM5 (BM1) are always positive (negative), so HF-PC3 is a value factor. Overall, the size factor is more important than the value factor at high frequencies.

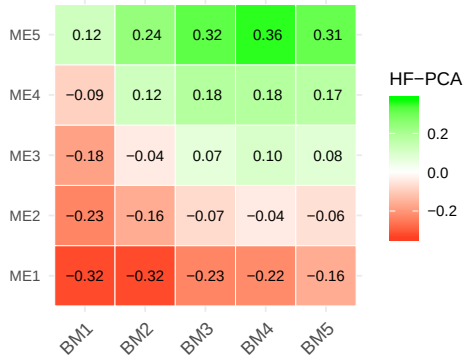
On the contrary, I observe the opposite at low frequencies. The heat-map in Panel (e) reveals that the second most crucial LF-PC is the value factor, while the size factor becomes the third-largest LF-PC, as is evident in Panel (f). This observation is largely compatible with the economic theory because the value factor often captures the business-cycle risk at low frequency. For example, Lettau and Ludvigson (2001b) point out that value stocks are more highly correlated with consumption growth in bad economic states than growth stocks, so they earn higher average returns. In short, the example in Figure 1 shows that the relative importance of latent factors can vary across frequencies.



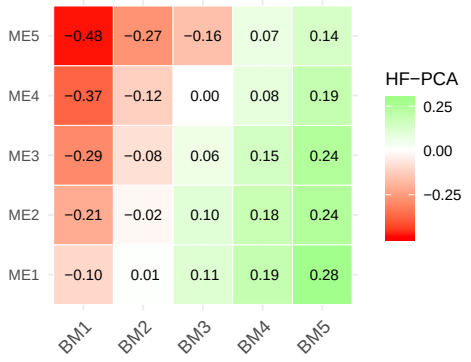
(a) 2nd PC of PCA



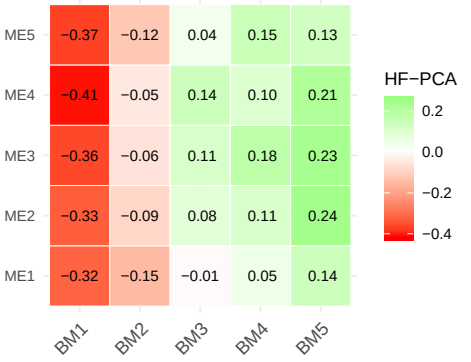
(b) 3rd PC of PCA



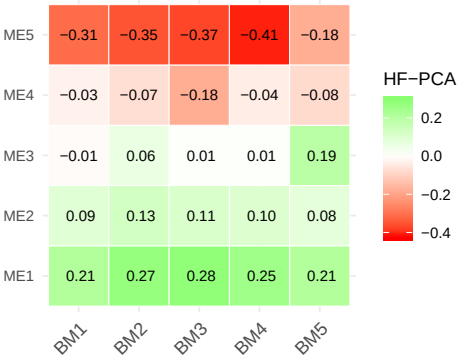
(c) 2nd PC of HF-PCA



(d) 3rd PC of HF-PCA



(e) 2nd PC of LF-PCA



(f) 3rd PC of LF-PCA

Figure 1: 25 Fama-French Size-Value Portfolios: 2nd and 3rd PCs

This figure shows the factor loadings of the second and third principal components in the cross-section of 25 Fama-French size-value portfolios. I estimate the factor loadings using the canonical PCA, HF-PCA, and LF-PCA. See Definition 1 for the algorithm. The sample is from August 1963 to December 2019.

III.3 Simple Simulation

In this section, I design a simple simulation to illustrate how the frequency-dependent PCA can recover the conditional information in asset returns. I assume that each asset return is driven by an IID systematic component (strong factor), a persistent state variable (weak factor), and an idiosyncratic element,

$$\boldsymbol{R_{t+1}} = \boldsymbol{\mu_R} + \boldsymbol{\beta_F} F_{t+1} + \boldsymbol{\beta_x} x_t + \boldsymbol{e_{t+1}}, \quad \boldsymbol{\beta_F^\top \beta_F} = 1, \quad \boldsymbol{\beta_x^\top \beta_x} = 1, \quad \boldsymbol{\beta_F^\top \beta_x} = 0. \tag{19}$$

$$F_{t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_F^2), \quad \boldsymbol{e_{t+1}} \overset{\text{iid}}{\sim} \mathcal{N}(\boldsymbol{0_N}, \sigma^2 \boldsymbol{I_N}), \quad F_{t+1} \perp x_t \perp \boldsymbol{e_{t+1}},$$
$$x_{t+1} = \rho_x \cdot x_t + \sqrt{1-\rho_x^2}\ \sigma_x\ \eta_{x,t+1}, \quad \eta_{x,t+1} \overset{\text{iid}}{\sim} \mathcal{N}(0,1). \tag{20}$$

In the above model, only the state variable x_t can predict asset returns, and it follows an AR(1) process. Examples of x_t include (1) the time-varying mean and variance of consumption growth in the long-run risk model (Bansal and Yaron (2004)), (2) the systemic- and stock-specific resilience in the recovery following a disaster in the disaster model (Gabaix (2012)), (3) the surplus consumption ratio in Campbell and Cochrane (1999), and (4) portfolio-level book-to-market ratio (Haddad, Kozak, and Santosh (2020)). Idiosyncratic shocks are homogeneous and have an identical variance σ^2.

According to example 4, the conditional information $\boldsymbol{\beta_x} x_t$ is asymptotically unidentified if $\sigma_x^2 < \sqrt{\lim \frac{N}{T}}\sigma^2$. Factor returns are weakly predicted, so the assumption for a relatively small σ_x^2 is reasonable. Furthermore, F_{t+1} and x_t are priced in the cross-section, so they enter the linear SDF,

$$\mathcal{M}_{t+1} = 1 - b_F \cdot F_{t+1} - b_x \cdot x_t, \tag{21}$$

where b_F and b_x are risk prices of F_{t+1} and x_t. The expected returns are determined by the fundamental asset pricing equation $\mathbf{E}[\mathcal{M}_{t+1}\boldsymbol{R_{t+1}}] = \boldsymbol{0_N}$, which implies

$$\boldsymbol{\mu_R} = -\text{Cov}(\mathcal{M}_{t+1}, \boldsymbol{R_{t+1}}) = b_F \boldsymbol{\beta_F} \sigma_F^2 + b_x \boldsymbol{\beta_x} \sigma_x^2. \tag{22}$$

F_{t+1} and x_t are latent, so I extract their tradable proxies. Specifically, I project them into the space of asset returns and find factor-mimicking portfolios with the highest Sharpe ratio:

$$F_{t+1} : \tilde{F}_{t+1} = \boldsymbol{\beta_F^\top R_{t+1}}, \quad \mathbf{E}[\tilde{F}_{t+1}] = b_F \sigma_F^2, \quad \text{Var}(\tilde{F}_{t+1}) = \sigma_F^2 + \sigma^2, SR_F^2 = \frac{b_F^2 \sigma_F^2}{1 + \frac{\sigma^2}{\sigma_F^2}},$$

$$x_t : \tilde{X}_{t+1} = \boldsymbol{\beta_x^\top R_{t+1}}, \quad \mathbf{E}[\tilde{X}_{t+1}] = b_x \sigma_x^2, \quad \text{Var}(\tilde{X}_{t+1}) = \sigma_x^2 + \sigma^2, SR_x^2 = \frac{b_x^2 \sigma_x^2}{1 + \frac{\sigma^2}{\sigma_x^2}},$$

where $\tilde{F}_{t+1}$ and $\tilde{X}_{t+1}$ are tradable proxies for F_{t+1} and x_t and are orthogonal by construction.

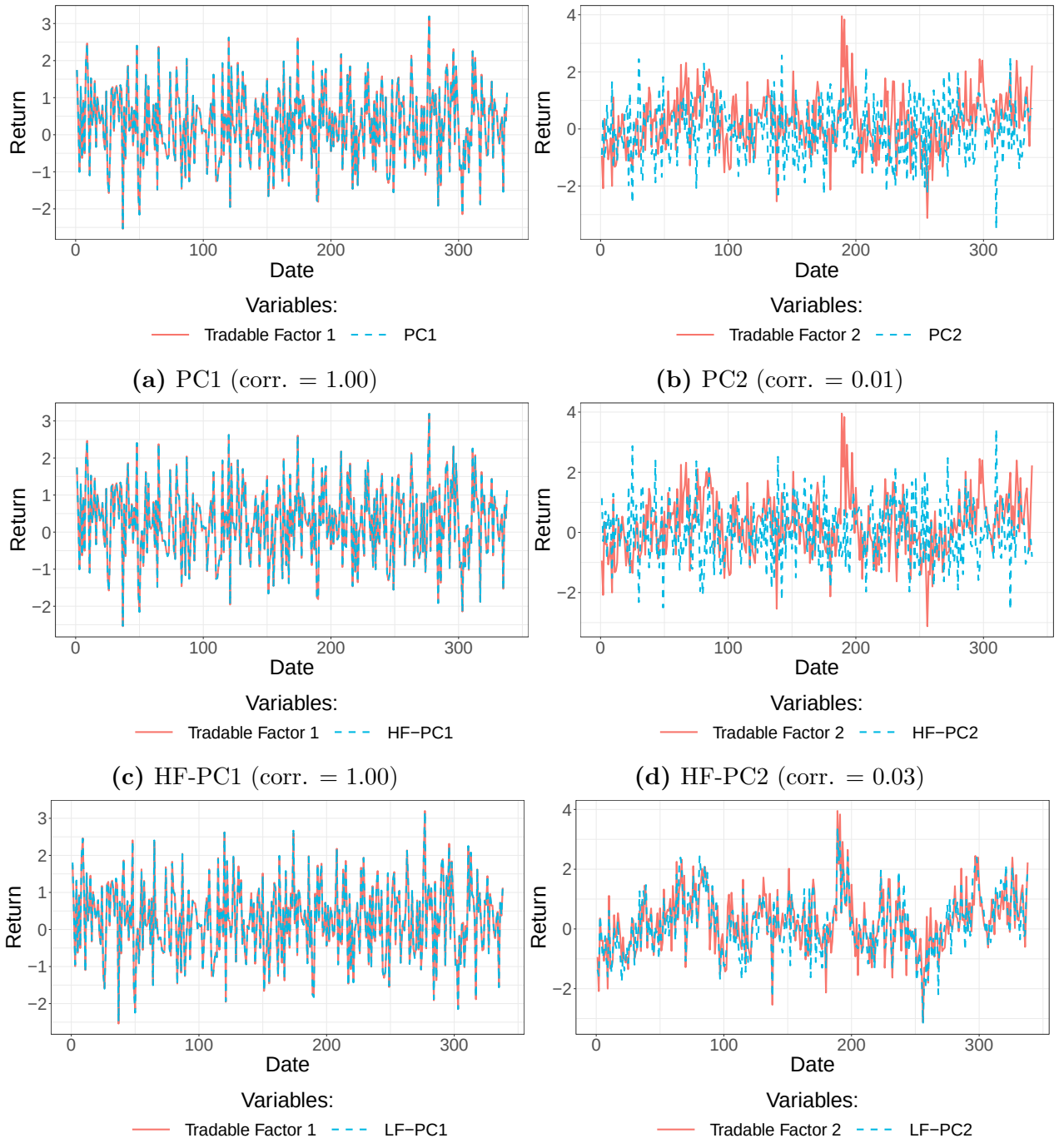
Regarding the simulation setup, I estimate the first PC in the cross-section of 78 test assets and assume that $\boldsymbol{\beta_F}$ is equal to its factor loadings and the volatility of F_{t+1} (σ_F) is 8.[12] Also, the idiosyncratic shocks have a unit variance ($\sigma^2 = 1$). In addition, the weak factor x_t has an identical factor loading ($\boldsymbol{\beta_x}$) as the second PC. However, its variance is small, $\sigma_x^2 = 0.5$; in other words, the weak factor explains a tiny fraction of time-series variations in single-period returns. In the cross-section of 78 test assets and 338 monthly observations, the canonical PCA has difficulty in identifying this factor: The critical value, $\sqrt{\lim \frac{N}{T}\sigma^2} \approx 0.48$, is close to the variance of weak factor. According to past literature, state variables that can predict asset returns tend to be extremely persistent, such as those in the long-run risk model, so I set ρ to be 0.9. Finally, I choose the Sharpe ratio of $\tilde{F}_{t+1}$ and $\tilde{X}_{t+1}$ to be 0.25 and 0.30 per month, and their risk prices can be reverse-engineered: $b_F = 0.032$ and $b_x = 0.735$.

Suppose I simulate factors and asset returns using the above model setup. I estimate the latent factors using canonical PCA, HF-PCA and LF-PCA. Figure 2 is one such example. The blue solid lines are "true" tradable factors $\boldsymbol{\beta_F^\top R_{t+1}}$ and $\boldsymbol{\beta_x^\top R_{t+1}}$, and the red dotted lines show the estimates of factors. As is evident in Panels (a), (c), and (e), I can always identify the first latent factor. Equivalently, the first latent factor explains the largest fraction of time-series variations in both short-horizon and long-horizon asset returns.

On the contrary, the weak factor is difficult to identify. Panels (b) and (d) show that neither canonical PCA and HF-PCA can recover the weak factor related to x_t, and the second PC or HF-PC has an almost zero correlation with the true factor-mimicking portfolio of x_t. The maximal Sharpe ratio implied by the first two PCs or HF-PCs is 0.276. However, the persistence of the state variables x_t magnifies its signal at low frequencies and allows me to detect it empirically. Panel (f) plots the second LF-PC, which closely tracks the "true" factor and has a correlation of 0.88. Moreover, the maximal Sharpe ratio implied by the first two LF-PCs is 0.321. Intuitively, as Figure A2 displays, the low-frequency variance of the persistent factor is much more considerable than its unconditional variance, so the signal of this factor passes the critical value $\sqrt{\lim \frac{N}{T}\sigma^2}$ at low frequencies and becomes identified.

Table 1 reports the simulation results of estimation using canonical PCA, HF-PCA, and LF-PCA in 1,000 simulations. The time-series sample size is 338. For each statistic, I show its 5th, 25th, 50th, 75th, 95th, mean, and mode. Panel (A) displays the correlation between the second true factor and estimated PC2 from canonical PCA, HF-PCA, and LF-PCA. Ideally, the correlation is 1. I focus on the second PC since the first PC is always identified, so there is no difference among different types of PCAs. The most important observation

[12] I assume a relatively large σ_F to make sure that it is a strong factor that is always identified by (HF- or LF- or canonical) PCA. The simulation results are robust to other $\sigma_F \in \{3, 5, 10\}$. In the data, the first three - five latent factors often have sizable eigenvalues (volatility) compared to the idiosyncratic volatility, so this assumption for the strong factor is reasonable.



(a) PC1 (corr. = 1.00)

(b) PC2 (corr. = 0.01)

(c) HF-PC1 (corr. = 1.00)

(d) HF-PC2 (corr. = 0.03)

(e) LF-PC1 (corr. = 1.00)

(f) LF-PC2 (corr. = 0.88)

Figure 2: Starting example: first two latent factors from canonical, HF-, and LF-PCA

I simulate one sample path of systematic factors and asset returns using the model setup described in the main text. This graph plots the time series of the first two "true" tradable factors (blue solid lines) and their estimates (red dotted lines) from canonical PCA (Panels (a) and (b)), HF-PCA (Panels (c) and (d)), and LF-PCA (Panels (e) and (f)). In addition, corr. refers to the correlation between the true tradable factor and its estimate. I standardize all time series to have unit variance.

is that the LF-PC2 has a much more significant correlation with the true second factor. Specifically, the average correlation between LF-PC2 and the true factor 2 is 0.754, whereas they are only 0.176 and 0.434 for HF-PC2 and canonical PC2, respectively. Hence, studying the LF components asset returns recovers a huge part of the persistent state variable.

In panel (B), I construct the MVE portfolios consisting of the first two latent factors: $\mathbf{MVE}_t = \hat{\boldsymbol{\mu}}_{\boldsymbol{F}}^{\top}\hat{\boldsymbol{\Sigma}}_{\boldsymbol{F}}^{-1}\boldsymbol{F_t}$, where $\boldsymbol{F_t}$ is either the first two PCs, HF-PCs, or LF-PCs. Since the LF-PCA recovers the persistent priced state variable, the LF-MVE portfolio has a greater Sharpe ratio than the other two portfolios. Panel (C) further reports the correlation among the HF-MVE, canonical-MVE, and LF-MVE portfolios. The MVE portfolios composed of the HF- and canonical PCs are highly correlated, with an average correlation of 0.937. However, the PCA can identify the state variable x_t in some simulations when the HF-PCA fails, so their MVE portfolios have correlations less than 0.746 in 5% of these simulations.

Finally, I decompose $\mathbf{MVE}_t^{HF}$ and $\mathbf{MVE}_t^{LF}$ in Panel (D),

$$\mathbf{MVE}_t^{LF} = \gamma^{HF}\mathbf{MVE}_t^{HF} + \mathbf{MVE}_t^{missing}, \quad \mathbf{MVE}_t^{HF} = \gamma^{HF}\mathbf{MVE}_t^{LF} + \mathbf{MVE}_t^{unpriced}.$$

I report the Sharpe ratio of $\mathbf{MVE}_t^{missing}$ and $\mathbf{MVE}_t^{unpriced}$ and also their correlation coefficients with the second factor. On average, the missing-MVE portfolio ($\mathbf{MVE}_t^{missing}$) has a correlation of 0.666 with the second true factor and yields a Sharpe ratio of 0.193. It implies that the HF-PCA misses important conditional information x_t. On the contrary, the unpriced MVE portfolio ($\mathbf{MVE}_t^{unpriced}$) has a negative correlation with the second true factor, but its Sharpe ratio is almost zero. Hence, the LF-MVE portfolio can be decomposed into two components: The first component, which is linear in the HF-MVE portfolio, identifies the IID shock driving a large proportion of common variations in asset returns, and the second component is the missing-part, mainly reflecting the slow-moving state variable.

In short, the simulation results confirm that both canonical PCA and HF-PCA often fail to identify the weak factor. However, if the weak factor is slow-moving, its signal can soar at low frequencies so that the LF-PCA can identify it.

III.4 Out-of-Sample Performance: 78 Test Assets

In this section, I examine a large cross-section of 39 firm characteristics from Kozak, Nagel, and Santosh (2020). Following the past literature such as Lettau and Pelger (2020b), I include both the short and long legs into my analysis, so there are 78 test assets. I focus only on two extreme portfolios for two reasons. First, if I consider all 10 sorted portfolios for each characteristic, there are 390 portfolios. This large cross-section is particularly challenging for the LF-PCA, which uses only long-run components of asset returns in estimation. It

Table 1: Simulation Results

	5th	25th	50th	75th	95th	Mean	Mode
Panel (A). Correlation between 2nd true factor and its estimate							
$corr(\boldsymbol{\beta_x^\top R_{t+1}}, (\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}^{\boldsymbol{PC}})^\top \boldsymbol{R_{t+1}})$	0.041	0.228	0.449	0.640	0.798	0.434	0.556
$corr(\boldsymbol{\beta_x^\top R_{t+1}}, (\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}^{\boldsymbol{HF}})^\top \boldsymbol{R_{t+1}})$	0.012	0.069	0.154	0.257	0.431	0.176	0.053
$corr(\boldsymbol{\beta_x^\top R_{t+1}}, (\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}^{\boldsymbol{LF}})^\top \boldsymbol{R_{t+1}})$	0.451	0.706	0.791	0.847	0.900	0.754	0.831
Panel (B). Sharpe ratio of MVE portfolios							
Sharpe ratio of $\mathbf{MVE}_t^{PC}$	0.179	0.232	0.272	0.313	0.374	0.274	0.264
Sharpe ratio of $\mathbf{MVE}_t^{HF}$	0.163	0.218	0.254	0.292	0.347	0.255	0.249
Sharpe ratio of $\mathbf{MVE}_t^{LF}$	0.210	0.279	0.325	0.378	0.473	0.330	0.332
Panel (C). Correlation between MVE portfolios							
$corr(\mathbf{MVE}_t^{PC}, \mathbf{MVE}_t^{HF})$	0.746	0.919	0.971	0.993	0.999	0.937	0.992
$corr(\mathbf{MVE}_t^{HF}, \mathbf{MVE}_t^{LF})$	0.497	0.665	0.789	0.899	0.979	0.770	0.917
$corr(\mathbf{MVE}_t^{PC}, \mathbf{MVE}_t^{LF})$	0.599	0.771	0.868	0.937	0.985	0.840	0.943
Panel (D). Difference between $\mathbf{MVE}_t^{HF}$ and $\mathbf{MVE}_t^{LF}$							
Sharpe ratio of $\mathbf{MVE}_t^{missing}$	0.035	0.114	0.189	0.260	0.363	0.193	0.199
Sharpe ratio of $\mathbf{MVE}_t^{unpriced}$	0.001	0.003	0.008	0.019	0.055	0.015	0.004
$corr(\boldsymbol{\beta_x^\top R_{t+1}}, \mathbf{MVE}_t^{missing})$	0.197	0.623	0.740	0.812	0.881	0.666	0.788
$corr(\boldsymbol{\beta_x^\top R_{t+1}}, \mathbf{MVE}_t^{unpriced})$	-0.741	-0.627	-0.514	-0.392	-0.087	-0.475	-0.520

This table reports the simulation results in 1,000 simulations. I estimate the systematic factors using canonical PCA, HF-PCA, and LF-PCA. For each statistic, I show its 5th, 25th, 50th, 75th, 95th, mean, and mode. In Panel (A), I consider the correlation between the second true factor and estimated PC2 from canonical PCA, HF-PCA, and LF-PCA. Ideally, the correlation should be 1. In Panel (B), I construct the mean-variance efficient (MVE) portfolios consisting of the first two latent factors: $\mathbf{MVE}_t = \hat{\boldsymbol{\mu}}_{\boldsymbol{F}}^\top \hat{\boldsymbol{\Sigma}}_{\boldsymbol{F}}^{-1} \boldsymbol{F_t}$, where $\boldsymbol{F_t}$ is either the first two canonical PCs, HF-PCs, or LF-PCs. Panel (c) reports the correlation between $\mathbf{MVE}_t^{PC}$, $\mathbf{MVE}_t^{HF}$, and $\mathbf{MVE}_t^{LF}$. Next, I decompose $\mathbf{MVE}_t^{HF}$ and $\mathbf{MVE}_t^{LF}$ in Panel (D) as follows:

$$\mathbf{MVE}_t^{LF} = \gamma^{HF} \mathbf{MVE}_t^{HF} + \mathbf{MVE}_t^{missing},$$

$$\mathbf{MVE}_t^{HF} = \gamma^{HF} \mathbf{MVE}_t^{LF} + \mathbf{MVE}_t^{unpriced}.$$

Finally, I report the Sharpe ratio of $\mathbf{MVE}_t^{missing}$ and $\mathbf{MVE}_t^{unpriced}$ and their correlation coefficients with the second factor.

implies a trade-off between signaling extraction and estimation noise, so I include only two extreme portfolios to control estimation errors. Also, when I include all 10 sorted portfolios, the portfolio weights are often the most enormous for portfolios in deciles 1 and 10. In other words, most of the relevant information comes from two extreme portfolios.

The entire sample is from August 1963 to December 2019. I further split the whole sample into two equal subsamples. Subsample 1 has 339 monthly observations, spanning from August 1963 to October 1991, and I treat it as the in-sample. Subsample 2 is the out-of-sample (OOS), which is from November 1991 to the end of the sample. As I show in Section II, estimating a linear SDF composed of asset returns is identical to finding the MVE portfolio

with the highest achievable Sharpe ratio. It requires me to focus on the OOS performance of asset pricing models, as the in-sample estimate often exaggerates the attainable Sharpe ratio in the real world. For instance, the annualized Sharpe ratio of 78 test assets is higher than 3 in the full sample, which is unreasonably large, according to the good deal bounds in Cochrane and Saa-Requejo (2000). In addition, McLean and Pontiff (2016) document the declining performance of many anomalies post-publication, and Kozak, Nagel, and Santosh (2018) also show that the average returns of 15 factors decrease considerably in the second subsample. Motivated by the previous papers, this paper estimates the PCs and their risk prices using data in the first subsample and evaluates the OOS performance in subsample 2.

HF vs. LF Time-Series Variations

First, I look at the time-series variations (TSVs) explained by different frequency-specific components in the in-sample. The results in the second subsample are largely similar (see Figure A4). I estimate the spectral density matrix $\hat{\boldsymbol{f}}_{\boldsymbol{R}}(\omega)$ via the DFT described in Appendix A.2, where the algorithm estimates only at frequencies $\frac{h}{360}$, $h \in \{1, \ldots, 180\}$. Specially, the HF component corresponds to the cycle length shorter than 36 months, that is, $\frac{360}{h} < 36$, whereas the LF part has a cycle period between 36 and 120 months, that is, $36 \leq \frac{360}{h} \leq 120$. Therefore, the sample estimates of $\boldsymbol{\Sigma}_{\boldsymbol{R}}^{HF}$ and $\boldsymbol{\Sigma}_{\boldsymbol{R}}^{LF}$ are as follows:

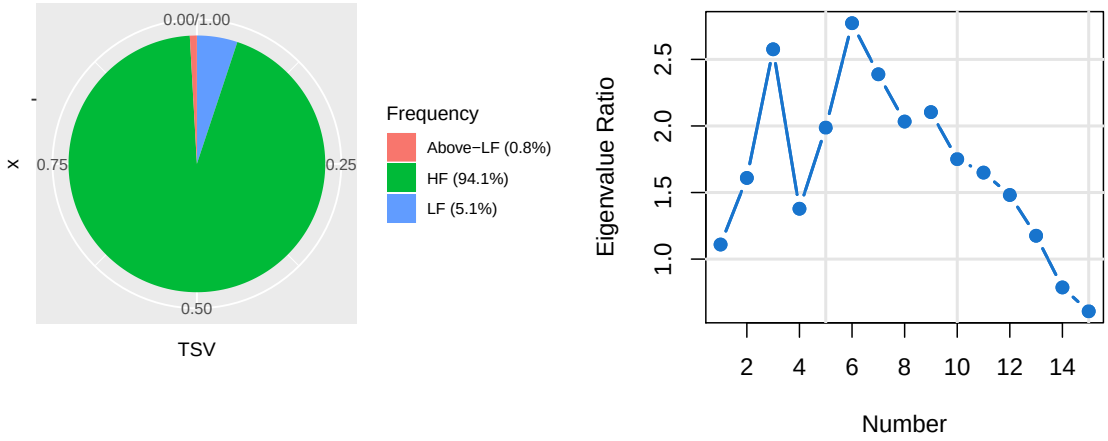
$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{HF} = \frac{1}{170}\sum_{h=11}^{180} \mathcal{R}\big(\hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big), \quad \hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{LF} = \frac{1}{8}\sum_{h=3}^{10} \mathcal{R}\big(\hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big).$$

I further define the fraction of time-series variations explained by HF and LF components as follows:

$$TSV^{HF} = \frac{tr\big[\sum_{h=11}^{180} \hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big]}{tr\big[\sum_{h=1}^{180} \hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big]} \quad \text{and} \quad TSV^{LF} = \frac{tr\big[\sum_{h=3}^{10} \hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big]}{tr\big[\sum_{h=1}^{180} \hat{\boldsymbol{f}}_{\boldsymbol{R}}(\frac{h}{360})\big]}.$$

If returns are uncorrelated, the spectral density matrix is approximately constant across frequencies, so the HF (LF, or above-LF) component accounts for $\frac{170}{180} = 94.5\%$ ($\frac{8}{180} = 4.4\%$, or $\frac{2}{180} = 1.1\%$) of time-series variations. Empirically, however, this LF part explains 5.1% of time-series variations, so this slow-moving component is slightly more important than that predicted by the uncorrelated assumption (see Figure 3(a)).

I further compare Tr $\big[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{LF}\big]$ to Tr $\big[\hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{HF}\big]$ and find that the former is around 1.2 times as the latter, which means the LF risk is slightly higher than the HF risk. Next, Figure 3(b) computes the ratio of LF-eigenvalues over HF-eigenvalues. An interesting observation is that the top 10 eigenvalues of the LF covariance matrix are 1.5 to 2.5 times as those of the HF one, except for the PC1. Therefore, the LF component has a clearer factor structure.



(a) Time-series variations explained by frequency-specific components

(b) LF-eigenvalues / HF-eigenvalues

Figure 3: Time-series variations in 78 assets, subsample 1

Panel (a) plots the fraction of time-series variations in 78 asset returns explained by the HF, LF, and above-LF components. Panel (b) plots the ratio of the first 15 low-frequency eigenvalues over high-frequency eigenvalues. The sample is from August 1963 to October 1991.

Out-of-Sample Sharpe Ratio

With the eigendecomposition of the frequency-dependent covariance matrix of asset returns, I construct OOS latent factors following definition 1: $\boldsymbol{F_t^{OOS}} = (\boldsymbol{Q}^{IN})^\top \boldsymbol{R_t^{OOS}}$, where $\boldsymbol{Q}^{IN}$ is the eigenvector of the frequency-dependent covariance matrix estimated in the first subsample, and $\boldsymbol{R_t^{OOS}}$ denotes asset returns in the out-of-sample. In addition, I estimate risk prices $\boldsymbol{b}$ for different prior Sharpe ratios in the in-sample, with the estimate denoted as $\hat{\boldsymbol{b}}^{IN}$. In the benchmark case, I use the objective function in equation (16) and include latent factors into the regression based on their eigenvalues. In other words, latent factors that drive more common movements among asset returns enter the SDF first. Next, I construct the OOS MVE portfolio, $\mathbf{MVE}_t^{OOS} = (\hat{\boldsymbol{b}}^{IN})^\top \boldsymbol{F_t^{OOS}}$.

Figure 4 is the heat-map of the OOS Sharpe ratio. I present only the HF- and LF-PCA results to save space, while the graphs of above-LF-PCA and PCA are in the appendix. In each panel, the x-axis denotes the prior Sharpe ratio of factor models, corresponding to different levels of L_2-shrinkage v_2 in equation (16). If I choose a larger prior Sharpe ratio, I impose a gentler shrinkage to risk prices $\boldsymbol{b}$. The y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. For example, the red color represents the "nearly" maximal monthly Sharpe ratio that these factor models

can achieve in the out-of-sample, around 0.35 - 0.38 in the data.

Panel (a) in Figure 4 and Panel (b) in Figure A5 show the results of HF-PCA and canonical PCA respectively — they have almost identical heat-maps. Generally, the first six or seven canonical or high-frequency PCs deliver an OOS Sharpe ratio of 0.28-0.29 across a wide range of prior Sharpe ratios. However, this low-dimensional (HF-)PC model still ignores an important priced component in the SDF. For example, when the prior Sharpe ratio is 0.4 or 0.5, in Figure 5, the OOS Sharpe ratio increases gradually from 0.28 to 0.37 as more (HF-)PCs enter the SDF. Besides, a substantial L_2-shrinkage helps reduce the required number of latent factors. Especially when $SR_{prior} = 0.2$, I need 20-25 (HF-)PCs to reach the nearly optimal OOS Sharpe ratio. The factor model composed of the extremely low-frequency PCs has a similar observation, as is indicated in Panel (a) in Figure A5. Since the above-LF-component is moving considerably slowly, estimating the covariance matrix is challenging, so I compare the HF and LF systematic factors and the SDFs composed of them in the following analysis.

Panel (b) plots the OOS Sharpe ratio of LF-PCs. A distinguishing feature is the sparsity in the space of LF-PCs. In Figure 5, the first *seven* LF-PCs are almost sufficient to achieve the optimal OOS Sharpe ratio, at around 0.37 per month. In other words, seven systematic factors that explain the most LF common variations in asset returns can span the whole asset space in the out-of-sample. Moreover, this observation is not sensitive to the choice of prior Sharpe ratio. With a wide range of reasonable prior, such as $SR_{prior} \in [0.3, 0.8]$, the SDF constructed by the first seven LF-PCs is always nearly optimal in the out-of-sample. Last but not least, since PCs are no more than linear transformations of original test asset returns $\boldsymbol{R_t}$, they contain almost identical information. Therefore, the MVE portfolios consisting of HF-PCs, LF-PCs, and original PCs earn just about the same OOS Sharpe ratio as the number of factors entering the SDF approximates N.

Out-of-Sample R^2_{gls}

In addition to the OOS Sharpe ratio, I also investigate the GLS R-squared of factor models, denoted by R^2_{gls}. With the in-sample estimate of risk prices, I compute the OOS pricing errors predicted by a factor model, $\boldsymbol{\alpha_R^{OOS}} = \bar{\boldsymbol{R}}_{\boldsymbol{t}}^{\boldsymbol{OOS}} - \text{Cov}(\boldsymbol{R_t^{OOS}}, \boldsymbol{F_t^{OOS}})\hat{\boldsymbol{b}}^{\boldsymbol{IN}}$, where $\text{Cov}(\boldsymbol{R_t^{OOS}}, \boldsymbol{F_t^{OOS}})$ is the sample covariance matrix between OOS asset returns $\boldsymbol{R_t^{OOS}}$ and OOS factors $\boldsymbol{f_t^{OOS}}$, $\bar{\boldsymbol{R}}_{\boldsymbol{t}}^{\boldsymbol{OOS}}$ is the sample average of OOS asset returns. R^2_{gls} is defined as

$$R^2_{gls} = 1 - \frac{(\boldsymbol{\alpha_R^{OOS}})^\top (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{\boldsymbol{OOS}})^{-1} \boldsymbol{\alpha_R^{OOS}}}{(\bar{\boldsymbol{R}}^{\boldsymbol{OOS}})^\top (\hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}^{\boldsymbol{OOS}})^{-1} \bar{\boldsymbol{R}}^{\boldsymbol{OOS}}}. \tag{23}$$

R^2_{gls} has a few satisfying properties. First, R^2_{gls} has a straightforward economic interpre-

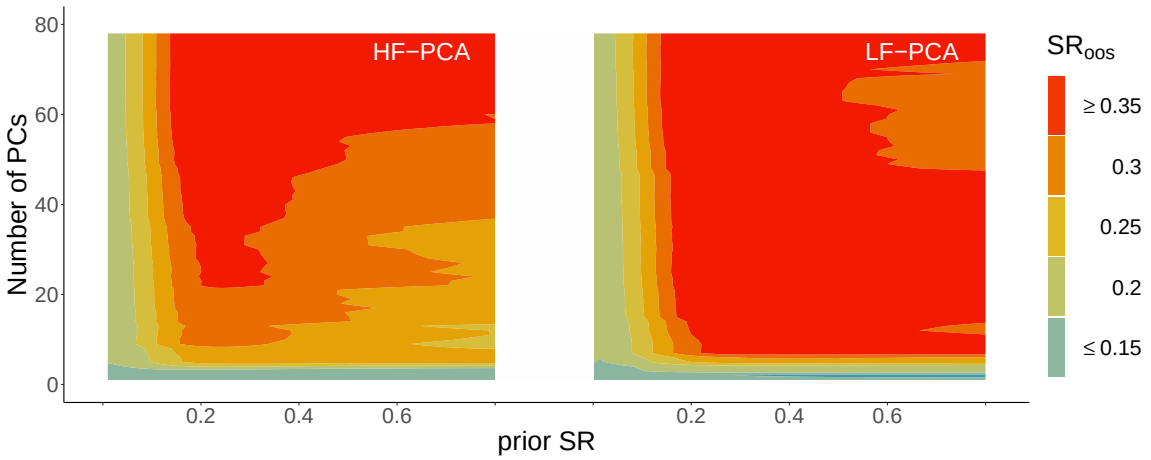
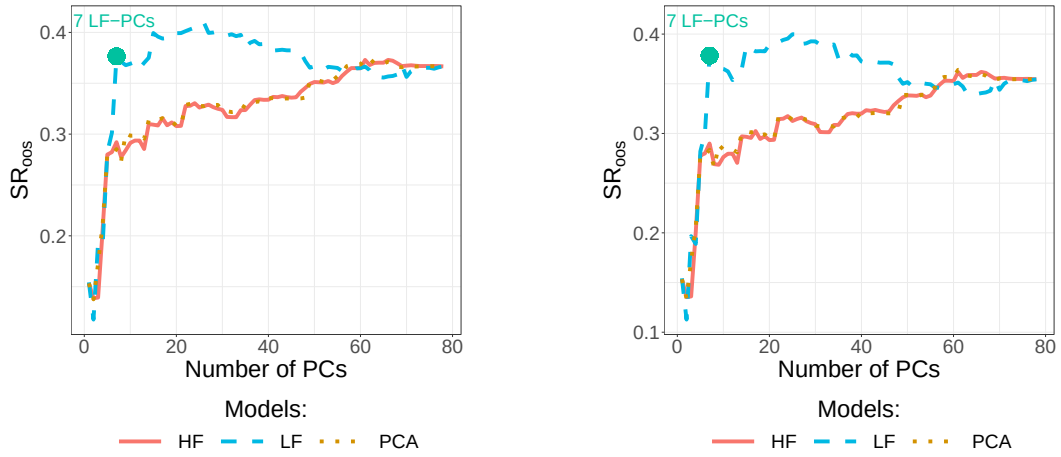


Figure 4: OOS Sharpe ratio of HF- vs. LF-PCA, 78 test assets

This graph plots the heat-maps of the OOS Sharpe ratio of HF- vs. LF-PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different levels of OOS Sharpe ratios. I include the PCs into the SDF based on their ability to explain time-series variations.



(a) $SR_{prior} = 0.4$ **(b)** $SR_{prior} = 0.5$

Figure 5: Zoom in OOS Sharpe ratio, $SR_{prior} \in \{0.4, 0.5\}$

This graph zooms in the OOS Sharpe ratio of PCA, HF-PCA, and LF-PCA. Different from figure 4, this figure shows the estimates using two prior Sharpe ratios, $SR_{prior} \in \{0.4, 0.5\}$.

tation: It quantifies the proportion of the squared Sharpe ratio of test assets explained by a factor model. Also, the objective function in equation (16) is to maximize R^2_{gls}. Therefore, it is natural to compare R^2_{gls} in the out-of-sample, consistent with my objective function. Last but not least, R^2_{gls} is invariant to any non-singular linear transformation of the original asset space. Specifically, for an arbitrary transformation of asset returns, such as $\boldsymbol{Y_t^{OOS}} = \boldsymbol{P}^\top \boldsymbol{R_t^{OOS}}$, where $\boldsymbol{P}$ is nonsingular, R^2_{gls} of pricing $\boldsymbol{Y_t^{OOS}}$ is exactly identical to that of $\boldsymbol{R_t^{OOS}}$. By focusing on R^2_{gls}, there is no need to choose whether the SDF should price original asset returns or their transformation, such as PCs.

Figure 6 plots the heat-maps of OOS R^2_{gls} of HF- and LF-PCA. Related plots of above-LF-PCA and original PCA can be found in Figure A6. Similar to the OOS Sharpe ratio, the PCA and HF-PCA share similar patterns — I need many latent factors to obtain the optimal OOS R^2_{gls}. On the contrary, I can choose a relatively parsimonious SDF consisting of LF-PCs. For instance, when SR_{prior} is between 0.5 and 0.8, the OOS R^2_{gls} of a 7 LF-factor-model is around 18% – 20%. Overall, the exploration of R^2_{gls} provides further evidence on the sparsity of LF-SDF. At the same time, the HF-SDF always needs more than 30 latent factors to achieve a nearly optimal OOS R^2_{gls}.

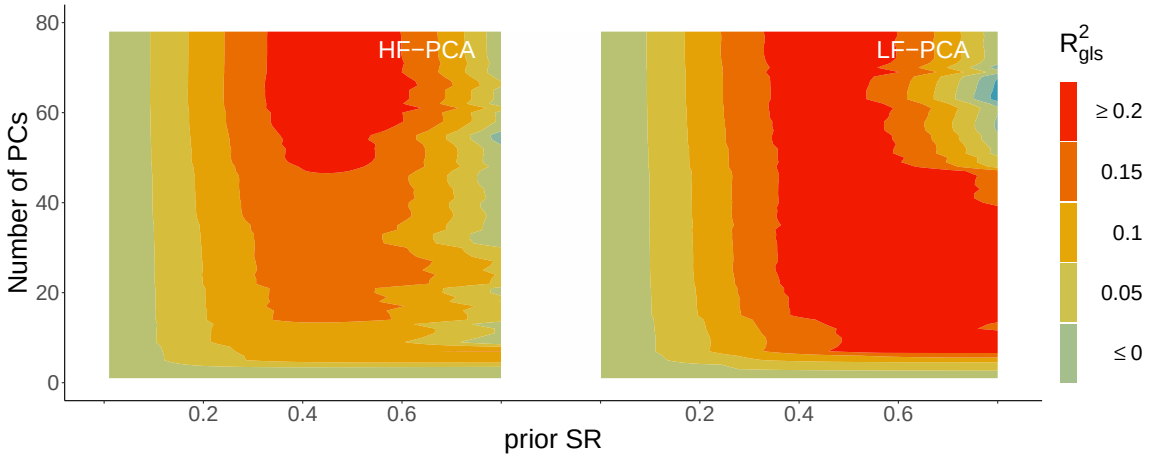


Figure 6: OOS R^2_{gls} of HF- vs. LF-PCA, 78 test assets

This graph plots the heat-maps of the OOS R^2_{gls} of HF- vs. LF-PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different levels of OOS R^2_{gls}. I include the PCs into the SDF based on their ability to explain time-series variations.

Zoom in High-Frequency Intervals

In the previous analysis, I define the HF interval as $\tau^{HF} \in [2, 36)$ and find the sparsity of latent factor models only at low frequencies. However, the definition of the HF interval is probably too wide to capture certain pricing information at a specific high frequency. For instance, the short-term reversal in Jegadeesh (1990) manipulates extremely fast-moving

information in predicting future stock returns. To explore whether the performance of latent factor models varies significantly under alternative definitions of HF intervals, I consider a further division of $\tau^{HF} \in [2, 36)$: (1) $[2, 3)$, (2) $[3, 6)$, (3) $[6, 12)$, and (4) $[12, 36)$. Next, I will examine the OOS Sharpe ratio of latent factor models in these four HF intervals.

Figure 7 plots the heat-maps of the OOS Sharpe ratio of latent factor models composed of PCs in these four HF intervals. Clearly, I always need more than 20 latent factors to achieve the optimal OOS Sharpe ratio. In addition, the performance of factor models is sensitive to the choice of the L_2-penalty — a significant penalty or a small prior Sharpe ratio is necessary to ensure a decent OOS performance. Hence, the sparsity of latent factor models only exists in the LF frequency interval $[36, 120)$.

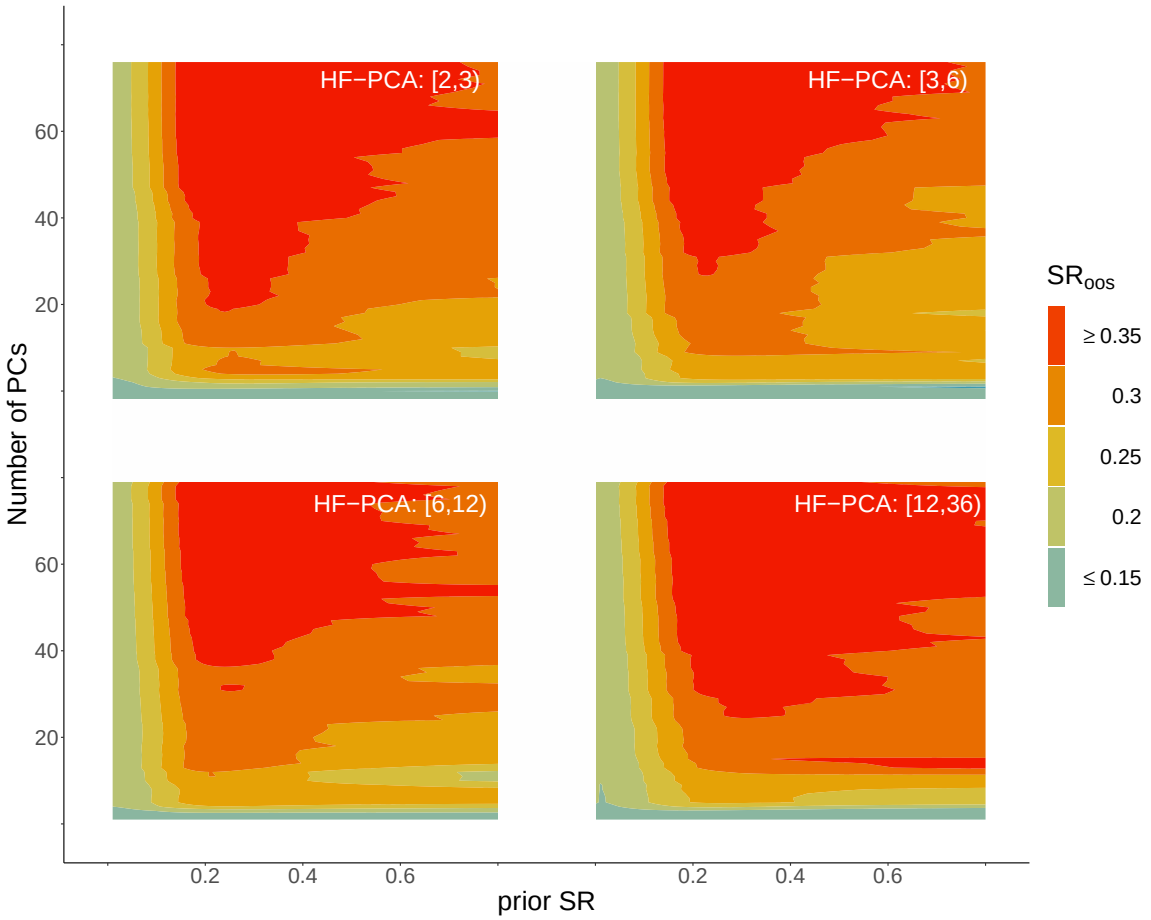


Figure 7: Robustness Check: OOS Sharpe ratio of 78 test assets, different HF intervals

This graph plots the heat-maps of OOS Sharpe ratios of HF-PCA, where I divide the HF intervals into four sub-intervals: $[2, 3)$, $[3, 6)$, $[6, 12)$ and $[12, 36)$. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF.

The previous empirical results shed light on the dynamics of priced state variables in the cross-section of 78 test assets. According to Proposition 2, the maximal Sharpe ratio implied by the SDF is frequency-dependent only due to state variables $\boldsymbol{X_t}$. More impor-

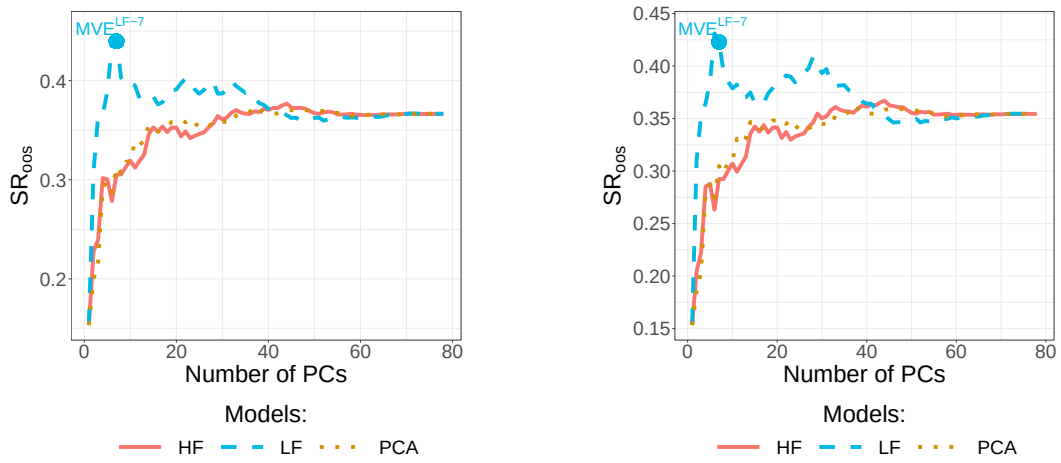
tantly, it implies that $\sum_{j=1}^{p} b_{X,i}^2 \boldsymbol{f}_{\boldsymbol{X_i}}(\omega)$, the second term in the spectral density function of the SDF, is on average larger at low frequencies. While either the canonical or HF-PCA fails to identify this persistent state variable, the LF-PCA recovers it as one of the largest factors explaining time-series variations in long-horizon asset returns. This conditional information is also priced in the cross-section. In the language of ICAPM (Merton (1973)), stock market participants have the incentive to hedge some slow-moving state variables, as those state variables can predict the future stock returns and economic environments and therefore affect investors' future investment opportunity set. Because of the hedging demand, the state variables command non-zero risk prices, so a valid SDF should not omit them. Finally, the fast-moving state variables are not essential in the monthly data. As Figure 7 indicates, the SDF is similarly dense in the space of extremely HF systematic factors.

Kozak, Nagel, and Santosh (2020) Estimation: Imposing Model Sparsity

As previous empirical results indicate, an SDF composed of (HF-)PCs cannot be parsimonious in terms of either the OOS Sharpe ratio or R_{gls}^2. What if I impose the sparsity of factor models? In this part, I follow the Kozak, Nagel, and Santosh (2020) procedure, described in equation (17), that includes L_1 shrinkage. According to the closed-form solution in equation (18), this procedure selects PCs with higher in-sample average returns first. Also, a larger v_1 renders more factors to have zero risk prices, so this algorithm enforces the sparsity of factor models.

I show the OOS Sharpe ratio of the MVE portfolios from the Kozak, Nagel, and Santosh (2020) estimation in Figure 8. First, 15 PCs or 20 HF-PCs and canonical PCs can deliver the nearly optimal Sharpe ratio, so the SDF does become sparser. However, it must be the case that the objective function chooses some small PCs that have essential pricing information. At the same time, the Kozak, Nagel, and Santosh (2020) procedure can still discover a sparse LF-SDF, with the first seven LF-PCs commanding a 0.4 monthly Sharpe ratio. According to the heat-maps of OOS Sharpe ratio in Figure A7, the LF-SDF is always sparse when the prior Sharpe ratio that I use to estimate the risk prices of LF latent factors is greater than 0.3.

Factor models that have been proposed in past literature are mostly sparse, such as the Fama-French three-factor model. However, there is no particular reason why a factor model must be sparse, even though people often pursue parsimonious models. Giannone, Lenza, and Primiceri (2021) call this "the illusion of sparsity." Moreover, it is almost unlikely to select a few firm characteristics, such as size and book-to-market ratio, and use them to span the whole asset space. For instance, Kozak, Nagel, and Santosh (2020) show that characteristics-sparse SDFs formed from a few factors cannot appropriately explain the cross-section of



(a) $SR_{prior} = 0.4$ **(b)** $SR_{prior} = 0.5$

Figure 8: OOS Sharpe ratio, Kozak, Nagel, and Santosh (2020) estimation

This graph zooms in the OOS Sharpe ratio of the canonical PCA, HF-PCA, and LF-PCA. The prior monthly Sharpe ratio is set to be 0.4 in Panel (a) and 0.5 in Panel (b). I estimates risk prices using the Kozak, Nagel, and Santosh (2020) algorithm described in equation (17). Latent factors are PCs of 78 test asset returns.

expected stock returns in the out-of-sample. In addition, Bryzgalova, Huang, and Julliard (2021) use a continuous spike-and-slab Bayesian prior to study 51 observable tradable and nontradable factors, and they find that within a wide range of reasonable prior Sharpe ratios of the SDF, 95% posterior credible intervals of the number of factors in the true model are between 16 and 32.

Why do we desire the sparsity of latent factor models? According to Kozak, Nagel, and Santosh (2018), the absence of near arbitrage opportunities implies that factors capturing the most systematic common variations in asset returns are non-diversifiable, so market participants earn non-zero risk premia for taking these risks. However, I observe some small (HF-)PCs bringing nontrivial risk premia. These factors explain less than 0.1% of the time-series variation, which implies that they are idiosyncratic shocks. Accordingly, they should not command sizable risk premia; otherwise, arbitrageurs can include those small PCs into their portfolio without increasing their investment risk significantly — instead, the sparsity of LF-SDF solves this puzzle to some extent.

How should we interpret the sparsity of the LF-SDF? On the one hand, it makes economic sense to observe a sparse LF-SDF. Suppose investors are buy-and-hold investors who pay more attention to the long-term trade-off between risk and returns, or investors have Epstein-Zin preference and are particularly risk-averse to the long-run uncertainty. Under these scenarios, the LF-SDF should imply a higher Sharpe ratio than the HF-SDF, because those LF systematic factors are the most risky in the long horizon. On the other hand, some persistent state variables explain a small fraction of common variations in single-period

returns, but they are much more prominent in the long horizon. Hence, the LF-PCA boosts the signal of this persistent conditional information and recovers them partially or wholly.

III.5 Do celebrated models explain HF and LF risks?

This section further compares the OOS MVE portfolios of latent factors with the following benchmark models: (1) CAPM, (2) Fama and French (1993) three factors (FF3), (3) Fama and French (2015) five factors (FF5), (4) Carhart (1997) four factors (Carhart4), and (5) Hou, Xue, and Zhang (2015) four factors (Q4). First of all, I examine whether these five sparse factor models can explain HF- and LF-MVE portfolios by running time-series regressions as follows:[13]

$$\mathbf{MVE}_t^Z = \alpha + \boldsymbol{\beta}^\top \boldsymbol{B_t} + \eta_t, \quad Z \in \{\text{HF}, \text{LF}\},$$

where $\boldsymbol{B_t}$ is one of the five benchmark models mentioned before. I report three test-statistics in Table 2: (1) α, (2) t-statistics of α, and (3) the adjusted R-squared, denoted as R^2_{adj}. To control for the serial dependence of pricing errors, I use Newey and West (1987) standard errors with both 36 lags (t-stat I) and 12 lags (t-stat II). In Table 2, I estimate risk prices of PCs under the prior Sharpe ratio equal to 0.4. To enhance interpretability, I normalize all MVE portfolios to have the same volatility as the market factor.

The first panel in Table 2 examines CAPM. Not surprisingly, the market factor alone entirely fails to explain neither HF- nor LF-MVE portfolios. The pricing errors are enormous, always greater than 1% per month. Moreover, LF-MVE portfolios always have higher alphas and t-statistics than HF ones; hence, they are more difficult ro explain. Interestingly, I observe relatively low R^2_{adj}, less than 10% in all columns. Since the first PC in the cross-section is always a level factor that is highly correlated with the market factor, the low R^2_{adj} in CAPM implies that the SDF loads heavily on other lower-order latent factors.

FF3 extends CAPM by including the size and value factors. Compared to CAPM, FF3 slightly reduces the pricing errors and significantly increases R^2_{adj}, particularly in the regression of HF-MVE portfolios. However, FF3 still fails to explain the OOS MVE portfolios of latent factors.

Carhart4 includes the momentum factor into FF3. Intriguingly, the alphas of MVE portfolios reduce by more than 40% compared to the previous two regressions, although all remain significantly positive. Moreover, the inclusion of the momentum factor improves the time-series fit dramatically. For example, Carhart4 explains 52% of time-series variation in the MVE portfolio of seven LF-PCs, while R^2_{gls} in FF3 is just 16%.

[13]Empirically, the MVE portfolios composed of the first several HF or canonical PCs are almost identical. Specifically, their correlation coefficients are around 98 – 99%. Therefore, I focus on comparing HF- and LF-MVE portfolios.

In the last two panels, I consider two models with both investment and profitability factors in them. Simply speaking, FF5 differs from Q4 in the additional value factor in FF5, and they adopt a slightly distinct approach to construct factors. In addition, Q4 is better at explaining the MVE portfolios than FF5. Notably, pricing errors of HF-MVE portfolios are remarkably smaller, declining to around 0.3% per month, and are no longer significant, except for t-statistic I in the column of seven HF-PCs . On the other hand, LF-MVE portfolios still have sizable and statistically significant pricing errors, at around 0.7% per month. I have similar empirical findings under another prior Sharpe ratio equal to 0.5 (see Table A3 in the appendix).

In short, none of the five benchmark models can explain LF-MVE portfolios, while the Q4 model in Hou, Xue, and Zhang (2015) is capable of rationalizing the abnormal returns of the HF-MVE portfolios.

Table 2: Do celebrated models explain HF and LF risks?

		Panel (A). $\mathbf{MVE}_t^{HF}$				Panel (B). $\mathbf{MVE}_t^{LF}$			
		7 PCs	8 PCs	9 PCs	10 PCs	7 PCs	8 PCs	9 PCs	10 PCs
CAPM	α	1.03%	1.00%	1.16%	1.18%	1.39%	1.38%	1.38%	1.41%
	t-stat I	(2.89)	(2.47)	(2.88)	(2.99)	(3.57)	(3.72)	(3.72)	(3.76)
	t-stat II	(3.05)	(2.72)	(3.16)	(3.28)	(4.48)	(4.60)	(4.60)	(4.60)
	R^2_{adj}	8.02%	5.55%	0.30%	0.44%	7.41%	6.38%	5.76%	4.36%
FF3	α	0.80%	0.77%	0.94%	0.98%	1.27%	1.27%	1.27%	1.29%
	t-stat I	(4.87)	(3.91)	(4.29)	(4.44)	(4.45)	(4.65)	(4.65)	(4.75)
	t-stat II	(4.76)	(4.13)	(4.60)	(4.73)	(5.82)	(5.95)	(5.97)	(6.05)
	R^2_{adj}	40.97%	38.39%	31.92%	26.44%	15.84%	14.24%	13.06%	12.05%
Carhart4	α	0.44%	0.39%	0.57%	0.59%	0.83%	0.82%	0.81%	0.84%
	t-stat I	(3.18)	(2.61)	(3.09)	(3.18)	(3.75)	(3.88)	(3.87)	(3.91)
	t-stat II	(2.81)	(2.39)	(3.08)	(3.14)	(4.40)	(4.46)	(4.45)	(4.44)
	R^2_{adj}	65.13%	65.62%	57.00%	53.91%	52.36%	51.2%	51.81%	50.21%
FF5	α	0.43%	0.43%	0.48%	0.53%	0.87%	0.84%	0.83%	0.84%
	t-stat I	(2.55)	(2.36)	(2.87)	(3.11)	(3.15)	(3.17)	(3.12)	(3.13)
	t-stat II	(2.43)	(2.42)	(3.24)	(3.43)	(3.94)	(3.88)	(3.81)	(3.87)
	R^2_{adj}	49.98%	48.09%	48.77%	41.91%	27.95%	28.02%	27.52%	27.26%
Q4	α	0.32%	0.23%	0.26%	0.33%	0.76%	0.73%	0.71%	0.72%
	t-stat I	(2.07)	(1.30)	(1.47)	(1.81)	(2.77)	(2.82)	(2.75)	(2.81)
	t-stat II	(1.60)	(1.10)	(1.30)	(1.61)	(3.25)	(3.22)	(3.14)	(3.24)
	R^2_{adj}	42.81%	39.87%	44.7%	40.47%	31.31%	31.35%	31.71%	31.16%

This table tests whether five sparse factor models proposed in past literature can explain the MVE portfolios composed of latent factors. I construct the MVE portfolios using the first seven to 10 latent factors following the same steps as in the section III.4. I estimate the factors' risk prices under the prior Sharpe ratio of 0.4. The five benchmark models include (1) CAPM, (2) Fama and French (1993) three factors (FF3), (3) Fama and French (2015) five factors (FF5), (4) Carhart (1997) four factors (Carhart4), and (5) Hou, Xue, and Zhang (2015) four factors (Q4). I report three test-statistic in table 2: (1) α, (2) t-statistics of α, and (3) adjusted R-squared, denoted as R^2_{adj}. To control for the serial dependence of pricing errors, I use Newey and West (1987) standard errors with both 36 lags (t-stat I) and 12 lags (t-stat II).

Next, I test whether LF-MVE portfolios can explain HF-MVE ones or whether the oppo-

site is valid. Similarly, I run time-series regressions, but the benchmark model $\boldsymbol{B_t}$ becomes either $\mathbf{MVE}_t^{HF}$ or $\mathbf{MVE}_t^{LF}$. Table 3 reports the results under the prior Sharpe ratio 0.4.

In Panel (a), I regress $\mathbf{MVE}_t^{HF}$ on $\mathbf{MVE}_t^{LF}$, both of which are constructed by the first seven, eight, nine, and 10 PCs. First, pricing errors are almost zeros in the statistical sense and less than 0.1% per month. In other words, the LF-MVE portfolios can span the HF-MVE portfolios. On the other hand, I regress $\mathbf{MVE}_t^{LF}$ on $\mathbf{MVE}_t^{HF}$ in Panel (b). Unlike Panel (a), pricing errors are always significantly positive, implying that the HF-MVE ignores an essential priced component of LF-MVE.

To sum up, the evidence in Tables 2 and 3 indicates that MVE portfolios, or SDFs, consisting of LF-PCs, should be the right benchmark. The first few LF-PCs can construct an LF-SDF that yields nearly optimal OOS Sharpe ratio, and none of the five notable factor models proposed in the past literature or HF-MVE portfolios can explain them. At the same time, they can fully explain HF-MVE portfolios in the out-of-sample.

Table 3: Which benchmark? HF vs. LF Tangency Portfolios

	Panel (A): $\mathbf{MVE}_t^{HF} = \alpha + \beta \mathbf{MVE}_t^{LF} + e_t$				Panel (B): $\mathbf{MVE}_t^{LF} = \alpha + \beta \mathbf{MVE}_t^{HF} + e_t$			
	7 PCs	8 PCs	9 PCs	10 PCs	7 PCs	8 PCs	9 PCs	10 PCs
α	-0.10%	-0.10%	0.10%	0.00%	0.60%	0.60%	0.70%	0.60%
t-stat I	(-0.74)	(-0.62)	(0.51)	(-0.08)	(3.29)	(4.19)	(2.79)	(2.70)
t-stat II	(-0.63)	(-0.52)	(0.44)	(-0.07)	(4.01)	(4.50)	(3.04)	(3.06)
R^2_{adj}	68.89%	62.68%	53.86%	63.23%	68.89%	62.68%	53.86%	63.23%

This table tests whether the LF-MVE portfolio can explain the HF-MVE or whether the opposite is valid. I construct the MVE portfolios using the first 7 – 10 latent factors following the same steps as in Section III.4. I estimate the factors' risk prices under the prior Sharpe ratio of 0.4. I report three test-statistic in Table 2: (1) α, (2) t-statistics of α, and (3) adjusted R-squared, denoted as R^2_{adj}. To control for the serial dependence of pricing errors, I use Newey and West (1987) standard errors with both 36 lags (t-stat I) and 12 lags (t-stat II).

III.6 Origins of Economic Risks in SDFs

Why do I observe the sparsity of latent factor models only in the space of LF-PCs? Why do sparse LF-MVE portfolios earn higher Sharpe ratios than those composed of the first few HF or canonical PCs? Do they represent different sources of economic fundamentals? This section attempts to answer these questions by studying the economic drivers behind the linear SDFs consisting of HF and LF systematic factors.

I consider the SDFs composed of the first *seven* HF-PCs or LF-PCs. I denote them as the HF-SDF and LF-SDF, respectively. From the heat-maps in Figure 4, the first seven LF-PCs can generate nearly optimal OOS Sharpe ratios under a wide range of prior distributions. In addition, the inclusion of extra PCs into the SDF adds enormous unpriced noises but

minimal additional pricing information. Last but not least, the space of the first seven HF-PCs is almost identical to that of the first seven canonical PCs, so I focus on comparing the HF-SDF to LF-SDF.

Past literature often uses the first several largest PCs of single-period returns, which are empirically identical to HF-PCs, to construct the linear SDF. However, my previous empirical findings indicate that such SDFs can neglect a vast priced component. Hence, I decompose the LF-SDF[14] ($\mathcal{M}_t^{LF}$) into two components, the first of which is perfectly correlated with the HF-SDF ($\mathcal{M}_t^{HF}$) and another of which is the orthogonal part as follows:

$$\mathcal{M}_t^{LF} = \beta_{HF}\mathcal{M}_t^{HF} + \mathcal{M}_t^{missing}, \quad \mathcal{M}_t^{HF} \perp \mathcal{M}_t^{missing}. \tag{24}$$

Similarly, I project the HF-SDF into the linear space of the LF-SDF and extract an uncorrelated component, denoted by $\mathcal{M}_t^{unpriced}$,

$$\mathcal{M}_t^{HF} = \beta_{LF}\mathcal{M}_t^{LF} + \mathcal{M}_t^{unpriced}, \quad \mathcal{M}_t^{LF} \perp \mathcal{M}_t^{unpriced}. \tag{25}$$

Table 4 reports the correlation matrix and Sharpe ratios implied by $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$. As mentioned before, $\mathcal{M}_t^{LF}$ implies a higher Sharpe ratio than $\mathcal{M}_t^{HF}$, and both SDFs imply statistically significant Sharpe ratios with t-statistics greater than 4. Moreover, $\mathcal{M}_t^{HF}$ accounts for only 69% of the time-series variation of $\mathcal{M}_t^{LF}$ but misses a considerable component $\mathcal{M}_t^{missing}$ that earns a monthly Sharpe ratio of around 0.2 and has a t-statistic equal to 4.6. In the following tables, I call $\mathcal{M}_t^{missing}$ the *missing-SDF*, which means that the traditional PCA or HF-PCA misses a huge priced component of the LF-SDF. Not surprisingly, the part of HF-SDF orthogonal to LF-SDF has almost zero Sharpe ratio, so this is an unpriced component. Hence, I will call $\mathcal{M}_t^{unpriced}$ the *unpriced-SDF*.

Table 4: Correlation among $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$

Corr.	$\mathcal{M}_t^{LF}$	$\mathcal{M}_t^{unpriced}$	$\mathcal{M}_t^{HF}$	$\mathcal{M}_t^{missing}$	SR	t-stat (36 lags)
$\mathcal{M}_t^{LF}$	1.00				0.376	6.22
$\mathcal{M}_t^{unpriced}$	0.00	1.00			0.037	0.67
$\mathcal{M}_t^{HF}$	0.83	0.56	1.00		0.292	4.71
$\mathcal{M}_t^{missing}$	0.56	-0.83	0.00	1.00	0.240	4.56

This table plots the correlation matrix and Sharpe ratio (SR) of the following four variables: $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$. $\mathcal{M}_t^{LF}$ and $\mathcal{M}_t^{HF}$ are OOS MVE portfolios composed of the first seven HF- or LF-PCs, and they are constructed by the procedures in Section III.4 under the prior Sharpe ratio 0.4. The last column reports the t-statistics of Sharpe ratio using the Newey and West (1987) standard errors with 36 lags. The out-of-sample period spans from November 1991 to December 2019.

[14]In this paper, the SDF is equal to one minus the MVE portfolio: $\mathcal{M}_t = 1 - \mathbf{MVE}_t$. Hence, it is equivalent to studying the MVE portfolios.

The findings in Table 4 also relate to literature that attempts to denoise the tradable factor. For instance, Golubov and Konstantinidi (2019) decompose the market-to-book ratio into market-to-value and value-to-book components — the market-to-value component drives nearly all the risk premium of the value strategy. In addition, Daniel, Mota, Rottke, and Santos (2020) document that unpriced components explain a reasonably large amount of Fama-French five factors, and they propose a novel way to hedge the unpriced components. By focusing on the long-term comovement of asset returns, the LF-SDF significantly reduces the unpriced component.

Dynamics of SDFs: Variance Ratio Test

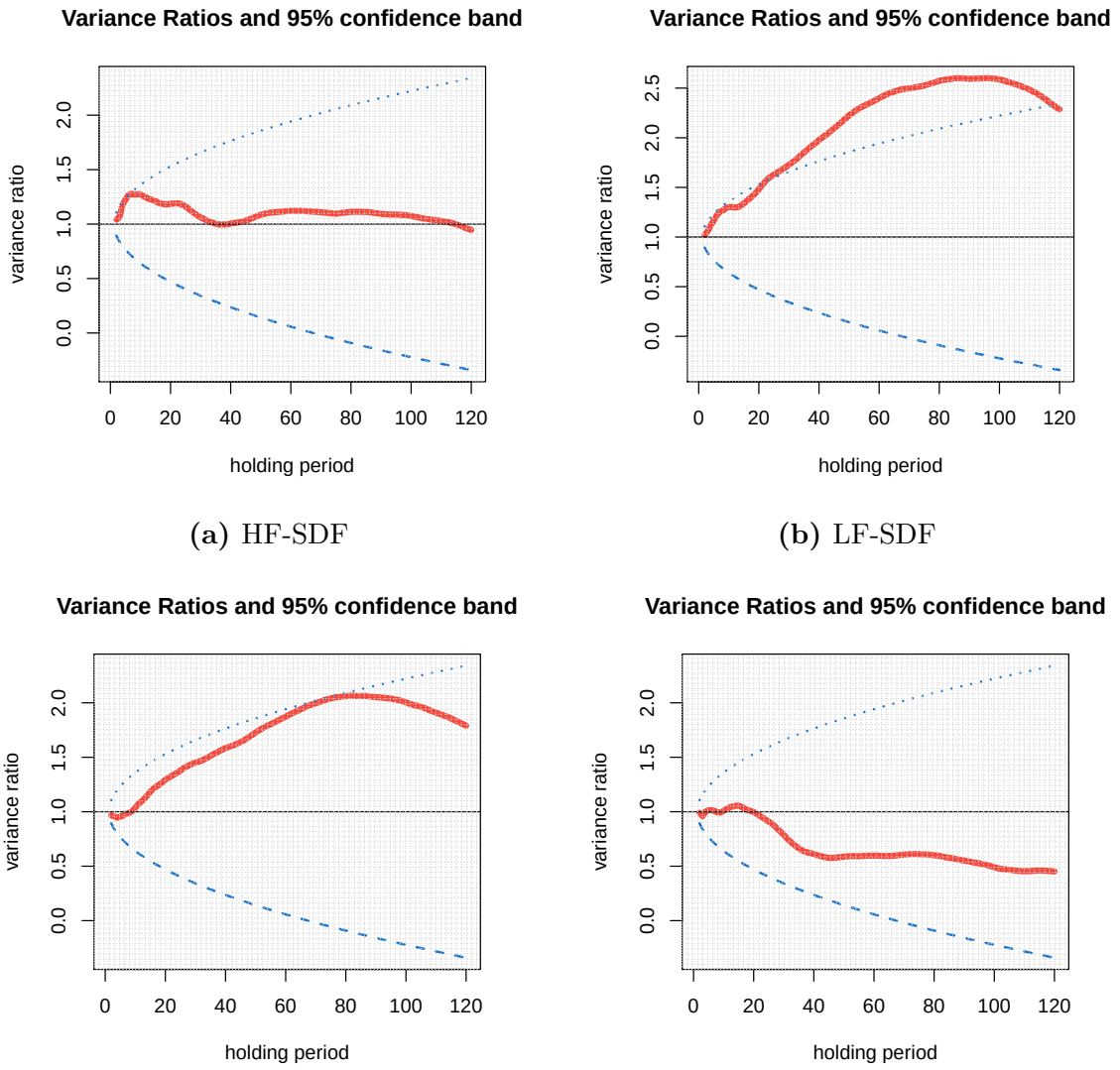
Theoretically, the LF-PCA has a better finite sample performance than the HF-PCA and canonical PCA because studying the long-horizon returns boosts the signal of some persistent conditional information driving the asset returns and detects them empirically. If the previous argument is valid, the LF-SDF must capture some conditional information ignored by the HF-SDF, so LF-SDF should have a different dynamic across multiple horizons. To explore the dynamics of SDFs, I resort to the variance ratio test, which is calculated as

$$VR(h) = \frac{\text{Var}(\mathcal{M}_{t,t+1} + \cdots + \mathcal{M}_{t+h-1,t+h})}{h \times \text{Var}(\mathcal{M}_{t,t+1})}, \tag{26}$$

where $\mathcal{M}_{t,t+1}$ is the single-period SDF. I can also rewrite the variance ratio as a weighted average of the autocorrelations of $\mathcal{M}_{t,t+1}$. A useful benchmark is the IID case, where the variance ratio test equals 1 at any horizon.

Figure 9 displays the variance ratios for $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$. The blue dotted lines are 95% confidence intervals of the variance ratios. If the solid red line crosses the dotted blue lines, I can reject the null hypothesis of the IID assumption. Panels (a) and (b) show the variance ratios for the HF-SDF and LF-SDF, respectively. While the HF-SDF exhibits limited autocorrelation over time, the LF-SDF displays a remarkable deviation from the IID assumption. For example, a five-year investor holding the LF-MVE portfolio is subject to double the variance of an investor with a monthly holding period. In addition, the variance ratio of the LF-SDF peaks between the six- and seven-year horizon, but it starts to decrease slowly after the seven-year horizon. Intuitively, the LF-SDF is riskier than HF-SDF from the perspective of long-term investors, so it should command a higher Sharpe ratio to compensate for bearing additional low-frequency risks.

Clearly, there are essential persistent components in the dynamic of the LF-SDF. Panel (c) further plots the variance ratio for the missing-SDF, which manifests a similar dynamic as the LF-SDF. Combining the evidence in Panels (a), (b), and (c), I conjecture that the



(a) HF-SDF

(b) LF-SDF

(c) Missing-SDF

(d) Unpriced-SDF

Figure 9: Variance Ratio of the SDF Components

This graph plots the variance ratios of $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$, calculated as

$$VR(h) = \frac{\text{Var}(\mathcal{M}_{t,t+1} + \cdots + \mathcal{M}_{t+h-1,t+h})}{h \times \text{Var}(\mathcal{M}_{t,t+1})},$$

where $\mathcal{M}_{t,t+1}$ is the single-period SDF. The HF-SDF and LF-SDF consist of the largest seven HF- and LF-PCs. The prior (monthly) Sharpe ratio used to estimated the risk prices is set to be 0.4. The blue dotted lines are 95% confidence intervals of the variance ratios. If the red solid line crosses the blue dotted lines, I can reject the null hypothesis of the IID assumption for the linear SDFs.

LF-SDF, which is the suitable benchmark and captures the highest attainable Sharpe ratio, contains two components: (1) the first component is spanned by the HF-SDF, which mainly captures the short-term information in asset returns and is roughly conditionally uncorrelated over time, and (2) the second component is ignored by canonical PCA, which identifies some persistent information that commands sizable risk premium. Panel (d) also presents the variance ratio for the unpriced-SDF. I confirm that this component mainly reflects short-horizon information, with a decreasing variance ratio after two years.

Cumulative Returns of MVE Portfolios

Next, I examine the cumulative performance from the perspective of an investor of the LF-MVE, HF-MVE, missing-MVE, and the market portfolio. To increase interpretability, I normalize all portfolio returns to have the same volatility as the market portfolio, about 4.2% per month. The portfolio weights of the MVE portfolios are determined by the data in the first subsample, so there is no looking-forward bias. Figure 10 plots the log of cumulative excess returns from November 1991 to December 2019 (OOS). The solid red line indicates that the LF-MVE portfolio has the best long-horizon performance, with a log cumulative excess return of around 5. As a comparison, the investor earns cumulative log-returns of 3.83 and 2.03 in the HF-MVE portfolio (solid blue line) and the market portfolio (solid green line), respectively. Another surprising fact is that investors of the MVE portfolios do not lose money during the dot-com bubble, while the market portfolio experiences a -40% return. However, the market values of all portfolios plummet during the 2008 global financial crisis.

The missing-MVE portfolio (solid orange line) is the component of the LF-MVE that is uncorrelated with the HF-MVE portfolio, and its behaviors are different from the HF-MVE portfolio. For instance, the missing-MVE portfolio has an extraordinary performance in the late 1990s, while the HF-MVE portfolio has an almost zero excess return during the same period.

With the decompositions of SDFs in equations (24) and (25), I examine how each component in Table 4 relates to economic risks. Specifically, I regress each economic variable on different SDF components and conduct statistical tests on the correlation coefficients between each economic variable and SDFs.

There are two primary objectives for these regressions. First, I attempt to understand the economics behind SDFs. For example, the LF-SDF implies a higher Sharpe ratio than the HF-SDF. However, the extra Sharpe ratio earned by the missing-SDF ($\mathcal{M}_t^{missing}$) is probably the compensation for bearing economic risks that the HF-SDF does not load on. Moreover, given the LF-SDF as the proper benchmark, I desire to learn whether different economic risks drive the HF component ($\beta_{HF}\mathcal{M}_t^{HF}$) and the persistent component ($\mathcal{M}_t^{missing}$).

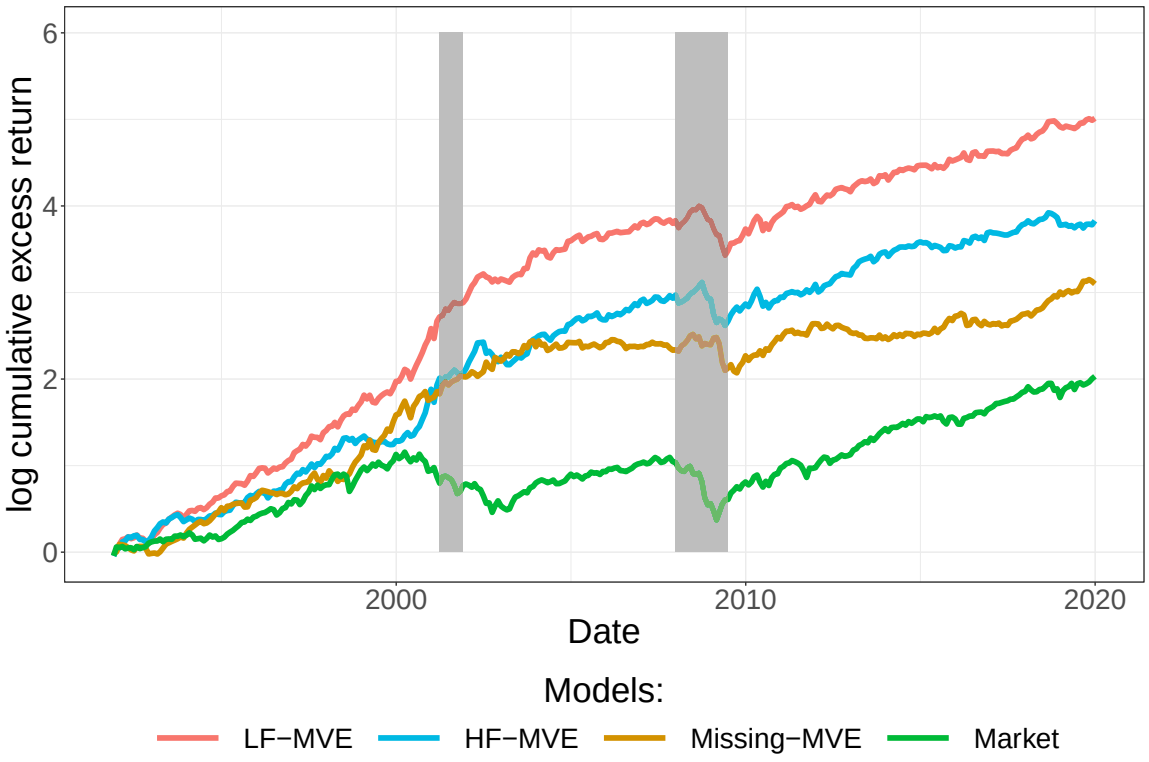


Figure 10: Log Cumulative Excess Returns of the MVE Portfolios

This graph plots the log of cumulative excess returns of the LF-MVE, HF-MVE, Missing-MVE (from table 4), and market portfolios. I normalize the LF-MVE, HF-MVE, Missing-MVE portfolios to have the same monthly volatility as the market portfolio. The sample spans from November 1991 to December 2019. Shaded areas denote the NBER recession periods: (1) 2001/03 – 2001/11 and (2) 2007/12 – 2009/06.

Second, it helps study the risk premium of a nontradable economic factor. As Cochrane (2009) indicates, we can define its risk premium as $-\text{Cov}(Y_t, \mathcal{M}_t)$, where Y_t is the nontradable factor. Similarly, Giglio and Xiu (2021) show that we should project a nontradable factor into the space of the largest principal components of a huge cross-section of test assets. However, Section III.4 points out that the SDF constructed by either PCs or HF-PCs potentially ignore an important priced component of the true SDF. According to equation (24), $\text{Cov}(Y_t, \mathcal{M}_t^{LF}) = \beta_{HF}\text{Cov}(Y_t, \mathcal{M}_t^{HF}) + \text{Cov}(Y_t, \mathcal{M}_t^{missing})$. An economic variable can be uncorrelated with $\mathcal{M}_t^{HF}$ but significantly correlated with the missing part, $\mathcal{M}_t^{missing}$. Hence, the study of LF latent factors provides additional insights into nontradable economic risks.

Table 5 reports the results. I consider eight economic variables, whose definitions are in Table A1. I standardize both dependent and independent variables so that readers can interpret all coefficient estimates as correlations. Similar to previous tables, I report two t-statistics using Newey and West (1987) standard errors with (1) 36 lags (t-stat I) and (2) 12 lags (t-stat II). Since macro variables are sometimes extremely persistent, I also report dependent variables' first-order autoregressive (AR(1)) coefficients (ρ). For example, if ρ is

close to 1, the economic variable is virtually a random walk process, making all statistical inference based on asymptotic normality invalid.

In Panel (A), I regress each nontradable economic variable Y_t on $\mathcal{M}_t^{HF}$ and $\mathcal{M}_t^{missing}$, while in Panel (B), I regress Y_t on $\mathcal{M}_t^{LF}$ and $\mathcal{M}_t^{unpriced}$. Since past literature often uses canonical PCs, which are almost identical to HF-PCs, it is intriguing to compare the coefficient estimates of $\mathcal{M}_t^{HF}$ and $\mathcal{M}_t^{LF}$. Also, if their coefficients are hugely different, the missing-SDF $\mathcal{M}_t^{missing}$ or the unpriced-SDF $\mathcal{M}_t^{unpriced}$ should explain the difference. In Table 5, I estimate the risk prices of factors under the prior Sharpe ratio 0.4. I consider a robustness check by adopting another prior Sharpe ratio 0.5, and Table A6 presents the related results. Overall, the results in Table 5 are not considerably different from those in A6.

Table 5: Economic Fundamentals related to HF- vs. LF-SDFs

Y_t :	C_t^{nd}	C_{t+1}^{nd}	GDP_t	GDP_{t+1}	N_t^{CF}	N_t^{DR}	HKM_t^{ntr}	HKM_t^{tr}	VXO_t^{ar1}	BW_t^{ar1}
Panel (A): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{HF} + \beta_2 \mathcal{M}_t^{missing} + \epsilon_t$										
$\mathcal{M}_t^{HF}$	-0.037	0.148	-0.175	0.037	-0.123	-0.299	-0.238	-0.293	0.238	-0.147
t-stat I	(-0.410)	(1.396)	(-0.761)	(0.347)	(-0.909)	(-2.926)	(-2.423)	(-2.484)	(2.525)	(-2.540)
t-stat II	(-0.412)	(1.396)	(-0.815)	(0.387)	(-0.805)	(-2.557)	(-2.315)	(-2.410)	(2.557)	(-2.716)
$\mathcal{M}_t^{missing}$	-0.218	-0.229	-0.180	-0.223	-0.112	-0.043	0.136	0.168	-0.040	-0.013
t-stat I	(-1.846)	(-3.953)	(-1.583)	(-2.203)	(-1.425)	(-0.610)	(1.467)	(1.581)	(-0.826)	(-0.171)
t-stat II	(-1.940)	(-3.953)	(-1.759)	(-2.484)	(-1.431)	(-0.591)	(1.460)	(1.572)	(-0.800)	(-0.171)
Panel (B): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{LF} + \beta_2 \mathcal{M}_t^{unpriced} + \epsilon_t$										
$\mathcal{M}_t^{LF}$	-0.147	0.004	-0.244	-0.088	-0.164	-0.272	-0.122	-0.149	0.175	-0.130
t-stat I	(-1.156)	(0.044)	(-1.003)	(-0.780)	(-1.425)	(-2.925)	(-1.420)	(-1.550)	(2.125)	(-1.922)
t-stat II	(-1.139)	(0.044)	(-1.063)	(-0.833)	(-1.426)	(-3.119)	(-1.564)	(-1.631)	(2.249)	(-1.921)
$\mathcal{M}_t^{unpriced}$	0.165	0.272	0.059	0.209	0.025	-0.131	-0.246	-0.303	0.166	-0.071
t-stat I	(3.023)	(3.513)	(0.733)	(2.300)	(0.201)	(-1.399)	(-2.328)	(-2.368)	(2.556)	(-1.074)
t-stat II	(2.634)	(3.410)	(0.939)	(2.683)	(0.195)	(-1.223)	(-2.140)	(-2.270)	(2.325)	(-1.081)
ρ	0.153	0.153	0.352	0.352	-0.189	-0.108	0.061	0.104	0.116	0.105
R^2_{adj}	4.91%	7.43%	6.30%	5.11%	2.76%	9.13%	7.51%	11.38%	5.80%	2.19%
Sample size	112	111	112	111	338	338	338	338	338	326

This table reports the results of the regressions in which I regress eight economic variables on different components of SDFs. The dependent variables include (1) and (2) current and one-period ahead quarterly real nondurable consumption growth, (3) and (4) current and one-period ahead quarterly real GDP growth, (5) cash-flow news, (6) discount-rate news, (7) nontradable intermediary factor, (8) tradable intermediary factor, (9) the AR(1) shock in VXO index, and (10) the AR(1) shock in investor (Baker and Wurgler (2006)) sentiments. The SDFs are composed of the first seven principal components of asset returns, and their risk prices are estimated under the prior Sharpe ratio equal to 0.4. I standardize both dependent and independent variables so that readers can interpret all coefficient estimates as correlations. I report two t-statistics using Newey and West (1987) standard errors with (1) 36 lags (t-stat I) and (2) 12 lags (t-stat II). In addition, I report dependent variables' first-order autocorrelation coefficients (ρ). The monthly (quarterly) out-of-sample runs from November 1991 to December 2019 (Q1 1992 – Q4 2019).

Quarterly Real Consumption Growth

First, I consider the textbook CCAPM, which predicts a negative correlation between

consumption growth and SDFs. However, past research (e.g., Kan and Zhang (1999)) find that the quarterly real nondurable consumption growth, commonly used in the past literature, is not strongly correlated with test assets. In other words, the risk premium of consumption risk is zero, contradicting the standard textbook prediction.

Column (1) in Table 5 presents the coefficient estimate of quarterly real consumption growth. Theoretically, when the consumption growth is low in bad economic states, marginal utility of investors, proxied by the SDF, should be higher, so economic theory predicts negative correlations. However, the correlation between consumption growth and the HF-SDF is only marginally negative, with a t-statistic of -0.4. However, Panel (B) shows that the consumption growth is closely associated with the LF-SDF, with a much higher correlation coefficient -0.15. The t-statistic (optimal lags) equals -1.2, so I cannot reject the null hypothesis of zero correlation.

What can explain this huge difference? The missing-SDF is the key, and its correlation with consumption growth is -0.22 and statistically significant. Suppose that a state variable predicts future consumption growth and portfolio returns, but it is relatively persistent. The standard PCA, which is virtually equivalent to HF-PCA, fails to identify the components related to this state variable; hence, the HF-SDF is not correlated with the consumption growth. However, the focus on the long-horizon asset returns recovers the identification of consumption risk. In short, the $\mathcal{M}_t^{missing}$ is significantly and negatively correlated with consumption growth, which implies a positive risk premium of consumption risk.

Campbell (1999) suggest an alternative timing convention to calculate the correlation between consumption growth and asset returns. Specifically, the consumption during a quarter is a flow. If we think of the consumption observed at quarter t as the consumption level at the beginning of this quarter, we should use the next-period consumption to compute consumption growth at quarter t. In other words, I should estimate the correlation between C_{t+1}^{nd} and $\mathcal{M}_t$. Column (2) in Table 5 displays the results. The missing-SDF is still significantly correlated with the next-period consumption growth, and its t-statistic is around -4.

Quarterly Real GDP Growth

Liew and Vassalou (2000) show that HML and SMB positively predict future real GDP growth. Motivated by this finding, I study whether the quarterly real GDP growth correlates with SDFs. Intriguingly, only the coefficient estimate of $\mathcal{M}_t^{missing}$ is significantly negative, with a t-statistic around -1.8 (see t-statistic II) in Column (3) of Table 5. Although the correlation coefficients of both $\mathcal{M}_t^{HF}$ and $\mathcal{M}_t^{LF}$ are not trivial, around -0.2, their standard errors are so enormous that I cannot reject the null hypothesis of zero correlation. The high autocorrelation coefficient, equal to 0.35, and small sample size, potentially contribute to

the notable estimation uncertainty.

In Column (4), I explore whether the SDFs can predict GDP growth in the next quarter. While both the HF-SDF and LF-SDF have almost zero prediction power, the missing-SDF negatively predicts the GDP growth. The coefficient estimate is -0.22 and has a t-statistic around -2. In other words, if the MVE portfolio implied by the missing-SDF experiences a negative return (or the missing-SDF increases) at quarter t, it predicts that the future GDP growth will decrease over the next quarter. This finding indicates that persistent state variables contained in asset returns can predict GDP growth. The missing-SDF captures this persistent predictor, so it is closely related to GDP growth.

Cash-Flow vs. Discount-Rate News

Campbell and Vuolteenaho (2004) decompose the shocks in the market portfolio into cash-flow news and discount-rate news.[15] In their language, cash-flow news is bad, for investors' wealth decreases and the future investment opportunity set is unchanged. On the contrary, discount-rate news is good since future investment opportunities, quantified by expected returns, improve.

Campbell and Vuolteenaho (2004) include four state variables: (1) the excess log return on the market, (2) the term yield spread that is the yield difference between 10-year and short-term constant-maturity taxable bonds, (3) the pricing-earnings ratio (PE) from Shiller (2000), and (4) the small-stock value spread that is the difference between $\log(\frac{BE}{ME})$ of the small high-book-to-market portfolio and $\log(\frac{BE}{ME})$ of the small low-book-to-market portfolio. However, the term yield spread that they used originally is no longer updated, so I replace it with the difference between the log yield on the 10-year U.S. Constant Maturity Bond and the log yield on the three-month U.S. Treasury bill, as in Campbell, Giglio, and Polk (2013). Campbell, Giglio, and Polk (2013) additionally include as a state variable the default spread (DEF), defined as the difference between the log yield on Moody's BAA and AAA bonds.

In the monthly data, I find that the default spread does not predict the market portfolio, so I stick to the four-state-variable VAR regression in Campbell and Vuolteenaho (2004). Moreover, I estimate the VAR model using monthly data from December 1928 to December 2019 and extract cash-flow and discount-rate news from November 1991 to December 2019. In Columns (5) and (6) of Table 5, I report the correlation coefficients between SDFs and two sources of shocks in the market portfolio.

[15] Campbell and Vuolteenaho (2004) estimate cash-flow and discount-rate news using a first-order VAR model: $\boldsymbol{Z_{t+1}} = \boldsymbol{a} + \boldsymbol{\Gamma Z_t} + \boldsymbol{u_{t+1}}$, where $\boldsymbol{Z_{t+1}}$ is a m-by-1 state vector with the log market excess return as its first entry. After estimating the VAR(1) model via OLS, they define cash-flow and discount-rate news as follows: $N_{t+1}^{CF} = \left[\boldsymbol{e_1}^\top + \boldsymbol{e_1}^\top \rho \boldsymbol{\Gamma}(\boldsymbol{I_m} - \rho \boldsymbol{\Gamma})^{-1}\right] \boldsymbol{u_{t+1}}$ and $N_{t+1}^{DR} = \boldsymbol{e_1}^\top \rho \boldsymbol{\Gamma}(\boldsymbol{I_m} - \rho \boldsymbol{\Gamma})^{-1} \boldsymbol{u_{t+1}}$, where $\boldsymbol{e_1}^\top = (1, 0, \ldots, 0)^\top$.

Cash-flow news is negatively correlated with $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{LF}$, but none of their coefficients is statistically significant. The LF-SDF is slightly more relevant to cash-flow news than the HF-SDF. Overall, the statistical power of these tests is not strong enough to make decisive conclusions.

Differently, discount-rate news is strongly and negatively correlated with both $\mathcal{M}_t^{HF}$ and $\mathcal{M}_t^{LF}$, with correlation coefficients around -0.3 and t-statistics around -3. In other words, discount-rate news earns a significantly positive risk premium. The time-series R^2, equal to 9% in column (6), is also considerably higher than in the regression of cash-flow news.

As a robustness check, I also estimate cash-flow and discount rate news by including the default spread as the fifth state variable in the VAR(1) regression. Columns (1) and (2) of Table A5 present similar results. The coefficient estimates in the regression are almost unchanged. In short, not only does discount-rate news explain most of the time-series variation in return news (see Campbell (1990)), but it is also more critical than cash-flow news as a source of economic risk for which investors in stock markets require risk compensation.

Intermediary Factor

He, Kelly, and Manela (2017) show that their intermediary factor can price many asset classes and conclude that financial intermediaries are important marginal investors and key to understanding asset prices. Their paper defines the intermediary capital ratio as the aggregate value of market equity divided by aggregate market equity plus aggregate book debt of primary dealers active. The intermediary capital risk factor, HKM_t^{ntr}, is the AR(1) innovation to the market-based capital ratio of primary dealers. He, Kelly, and Manela (2017) also define a tradable intermediary factor, denoted as HKM_t^{tr}. As predicted by intermediary asset pricing theory, such as He and Krishnamurthy (2013), the SDF of financial intermediaries is higher when a negative shock hits them, so the correlation between their SDF and the intermediary factor is expected to be negative.

Column (7) in Table 5 studies the nontradable intermediary factor. Panel (A) shows that the HF-SDF has a significant negative correlation with HKM_t^{ntr}, equal to around -0.24 with a t-statistic of -2.4. However, the LF-SDF has a smaller correlation (-0.12) in absolute terms, and its t-statistic is only -1.4. Hence, I am on the edge of rejecting the null hypothesis of zero correlation between HKM_t^{ntr} and $\mathcal{M}_t^{LF}$. More surprisingly, $\mathcal{M}_t^{missing}$ is positively associated with the intermediary factor, which implies that it hedges the intermediary risk in the HF-SDF. In other words, the intermediary factor cannot explain the high Sharpe ratio of $\mathcal{M}_t^{missing}$, or makes it even more puzzling.

Column (8) in Table 5 runs similar regressions but uses the tradable intermediary factor. The observations are largely compatible with those in Column (7). I also report the correla-

tion between SDFs and quarterly intermediary factors in Columns (3) and (4) of Table A5, and the empirical patterns are virtually identical.

Even though I do not discover a significant correlation coefficient between the LF-SDF and the intermediary factor, it does not imply that financial intermediaries do not play an important role in understanding asset prices. On the one hand, the intermediary factor is significantly correlated with the HF-SDF, especially its unpriced component, so the intermediary factor, at the very least, drives the common variations in asset returns. On the other hand, the risk premium of the nontradable factor in He, Kelly, and Manela (2017) is not statistically different from zero in the monthly regression of stock portfolios, which is consistent with the insignificant correlation between the LF-SDF and the intermediary factor. Also, financial intermediaries should be more important in other asset markets, such as CDS and derivative markets, in which they get involved actively.

Jump Risk

Investors require compensation for bearing downside risk (e.g., Ang, Chen, and Xing (2006)). While there is no consensus on which variable represents downside risk, I use the VXO index as the proxy, since it is commonly accepted as the fear index in the industry. Specifically, the VXO index is the risk-neutral entropy of the market excess return and is particularly sensitive to the left tail of the return distribution.

It is problematic to regress the VXO index on SDFs. The VXO index is highly persistent, with an AR(1) coefficient of around 0.9, so standard errors of coefficient estimates are enormous. In other words, the high persistence makes the statistical inference almost impossible in small samples. Hence, I extract the shock in the VXO index via an AR(1) regression, $Y_t = a + \rho Y_{t-1} + n_t$, and jump risk is defined as the AR(1) innovation $Y_t - a - \rho Y_{t-1}$.

Column (9) of Table 5 reports the correlation between SDFs and jump risk. Both the HF-SDF and LF-SDF have significantly positive correlations $(0.18 - 0.24)$ with jump risk. Interestingly, coefficient estimates of the HF-SDF and LF-SDF in Table 5 are similar to those in Table A5, in which I regress the original VXO index on SDFs. Hence, it can increase the power of statistical tests to focus on the much less persistent AR(1) innovation.

Intuitively, when investors are particularly fearful, the SDFs, proxying for their marginal utility functions, are likewise high. In other words, investors are willing to pay a positive risk premium to hedge jump risk. Nevertheless, the missing-SDF is almost unrelated to jump risk, so jump risk does not explain the risk premium of $\mathcal{M}_t^{missing}$.

Investor Sentiment

Rational economic models cannot always explain economic phenomena that we observe

in the real world, such as the tech stock bubble in the late 1990s and the housing bubble in 2008. Instead, investor sentiments are also essential in understanding asset prices. For instance, De Long, Shleifer, Summers, and Waldmann (1990) build a theoretical model in which the presence of noise traders with stochastic beliefs can create a source of risk that requires a positive risk premium. Kozak, Nagel, and Santosh (2018) show that if the demand from sentiment investors drives a large proportion of asset returns' common variations, their demand shocks, or investor sentiments, should enter the SDF as well.

Motivated by these papers, I go on to explore how SDFs extracted purely from asset returns correlate with the proxy for investor sentiments. First of all, I use the BW sentiment index in Baker and Wurgler (2006), which estimate the first principal component of six variables: the closed-end fund discount, the NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium.

The AR(1) coefficient of the BW index is close to 1, so I extract its AR(1) shock following the same steps as for the VXO index. The last column of Table 5 demonstrates that the HF-SDF is negatively correlated with the investor sentiment, with a t-statistic of around -2.6. The LF-SDF has a similar coefficient estimate (-0.13), and its t-statistic is about -1.9. Column (6) of Table A5 reports the correlation between SDFs and the original BW sentiment. Even though the magnitudes of coefficient estimates are extremely similar, their t-statistics are much lower due to the persistence of the BW sentiment index. Last but not least, the missing-SDF is virtually unrelated to the BW sentiment index. Overall, Table 5 indicates that only macro risk, such as consumption and GDP growth, can potentially explain the risk premium of $\mathcal{M}_t^{missing}$.

Huang, Jiang, Tu, and Zhou (2015) modify the BW sentiment index using the partial least squares method. Precisely, they extract the most important component that can simultaneously predict the future market return and explain time-series variations of the original six proxies. I call their sentiment index HJTZ sentiment. Columns (7) and (8) of Table A5 show that only the HF-SDF is weakly correlated with the AR(1) shock of HJTZ sentiment. On the contrary, its correlation with the LF-SDF is only -0.07, compared to -0.14 for BW sentiment. Overall, the HJTZ sentiment is less correlated with SDFs than the BW sentiment.

Summary

The findings in Table 5 deepen our understanding of the economics behind the factor zoo. One potential reason for the existence of the factor zoo is that factors are noisy proxies for economic fundamentals and therefore do not span each other. For example, Liew and Vassalou (2000) report that both the value and size factors can predict GDP growth, but they are never comprehensive predictors and cannot replace each other. This paper further

shows that the importance of economic risks varies at different frequencies. In Table 5, I confirm that a sparse LF-SDF, earning a nearly optimal Sharpe ratio, captures two elements: (1) the first one is perfectly linear in the HF-SDF and almost uncorrelated over time, which is statistically associated with discount-rate news of the market excess return, the intermediary factor, jump risk, and investor sentiment, whereas (2) the second one is neglected by the HF-SDF and captures some persistent state variables, reflecting business-cycle risks related to consumption and GDP growth.

IV Additional Robustness Check

In this part, I present a few of the robustness checks of Section III. Specifically, I investigate whether the sparsity of the LF-SDF is robust (1) when I consider only long-short portfolios, (2) if I impose the CAPM, or (3) if I slightly modify the definition of the LF interval.

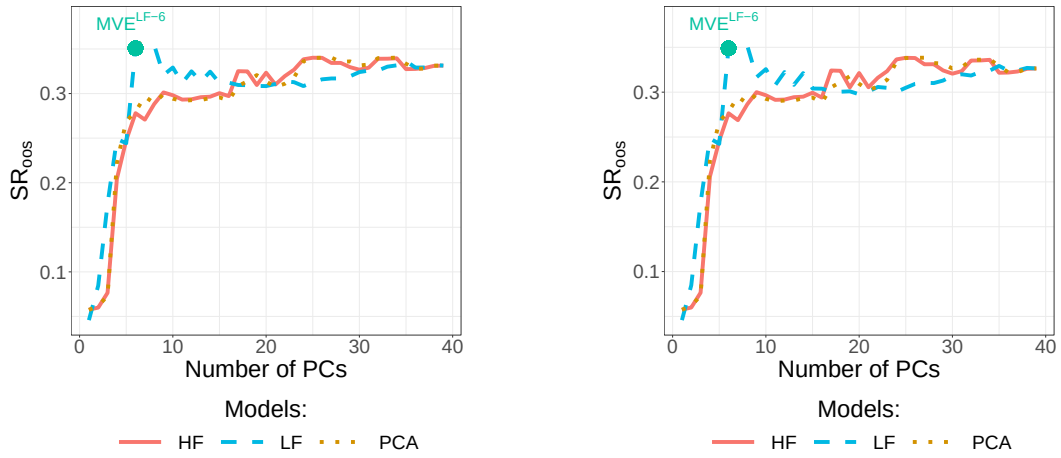
IV.1 39 Long-Short Portfolios

Until now, I have included long and short portfolios separately for each firm characteristic in Section III. However, many papers in cross-sectional asset pricing literature handle long-short portfolios, such as the size factor in FF3. To confirm the robustness of the main results, I further analyze long-short portfolios of 39 firm characteristics.

Figure 11 plots the OOS Sharpe ratio of PCA, HF-PCA and LF-PCA under prior Sharpe ratios $\in \{0.4, 0.5\}$. A more comprehensive heat-map is in Figure A8. First, the maximal Sharpe ratio is around 0.35, slightly less than that in the cross-section of 78 portfolios. Second, I can still discern a parsimonious factor model composed of low-frequency PCs. Particularly, a six-factor LF-PC model delivers an optimal OOS Sharpe ratio, and this finding is robust across a wide range of L_2-penalty, as I observe in Figure A8. On the contrary, latent-factor models constructed by canonical and high-frequency PCs are dense, consistent with my observations in Section III.4. Overall, the sparsity of LF-PC models is robust in the cross-section of 39 long-short portfolios. In the following robustness analyses, I stick to the original cross-section of 78 test assets.

IV.2 Imposing the CAPM

Since the introduction of the CAPM by Sharpe (1964) and Lintner (1965), the market factor has become the most influential factor in cross-sectional asset pricing. For example, Barillas and Shanken (2018) use the market factor as the anchor to compare a few famous factor models via Bayes factors. Also, Kozak, Nagel, and Santosh (2020) extract the CAPM α of



(a) $SR_{prior} = 0.4$ **(b)** $SR_{prior} = 0.5$

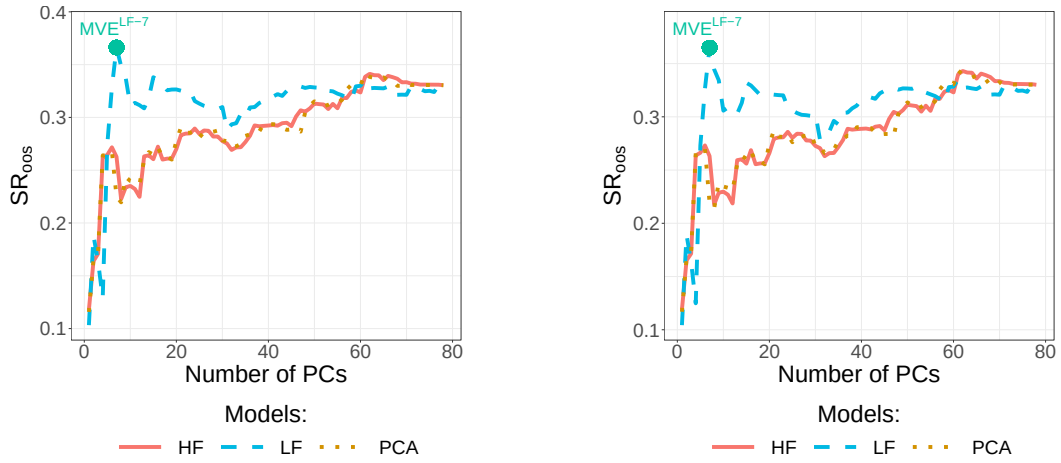
Figure 11: Zoom in OOS Sharpe ratio of 39 long-short portfolios, $SR_{prior} \in \{0.4, 0.5\}$

This graph zooms in the OOS Sharpe ratio of PCA, HF-PCA, and LF-PCA. The cross-section of test assets is 39 long-short portfolios. Different from figure A8, this figure shows the estimates using two prior Sharpe ratios, $SR_{prior} \in \{0.4, 0.5\}$.

50 long-short anomalies and estimate other systematic factors via the eigendecomposition of the CAPM $\boldsymbol{\alpha}$. Following the past literature, I turn to study the CAPM $\boldsymbol{\alpha}$.

Here is my empirical strategy. First, I regress $\boldsymbol{R_t}$ on the market factor using the in-sample observations: $\boldsymbol{R_t^{IN}} = \boldsymbol{\beta_0^{IN}} + \boldsymbol{\beta_m^{IN}} R_{mt}^{IN} + \boldsymbol{e_t^{IN}}$, and the CAPM $\boldsymbol{\alpha}$ is defined as: $\boldsymbol{\alpha_t^{IN}} = \boldsymbol{R_t^{IN}} - \boldsymbol{\beta_m^{IN}} R_{mt}^{IN}$. Next, I decompose the covariance matrix of $\boldsymbol{\alpha_t^{IN}}$ into frequency-dependent components and estimate frequency-dependent PCs as in definition 1. When I mention a K-factor model, the SDF consists of the market factor and the first K PCs of $\boldsymbol{\alpha_t^{IN}}$: $\mathcal{M}_t = 1 - b_m(R_{mt}^{IN} - \mu_m) - \boldsymbol{b_F}^\top(\boldsymbol{F_t} - \boldsymbol{\mu_F})$. Finally, I estimate risk prices $(b_m, \boldsymbol{b_F}^\top)^\top$ using the objective function in equation (16). To evaluate the OOS performance, I use the in-sample estimate of market loadings $\boldsymbol{\beta^{IN}}$ to construct the OOS CAPM $\boldsymbol{\alpha}$: $\boldsymbol{\alpha_t^{OOS}} = \boldsymbol{R_t^{OOS}} - \boldsymbol{\beta^{IN}} R_{mt}^{OOS}$. Then I construct the OOS latent factors and MVE portfolio as before.

Figure 12 plots the Sharpe ratio of the OOS MVE portfolio following the procedures described in the previous paragraph. The prior monthly Sharpe ratio is set to be 0.4 in Panel (a) and 0.5 in Panel (b). Like the benchmark case, the MVE portfolio consisting of the market factor and another 6 LF-PCs can earn a virtually optimal OOS Sharpe ratio, around 0.37 monthly. HF-PCA, however, needs much more PCs, literally more than 60, to reach the highest point, and PCA has almost an identical pattern. In short, a seven LF factor model can nearly span the whole asset space in the out-of-sample.



(a) $SR_{prior} = 0.4$ **(b)** $SR_{prior} = 0.5$

Figure 12: Imposing CAPM: OOS Sharpe ratio of 78 portfolios, $SR_{prior} \in \{0.4, 0.5\}$

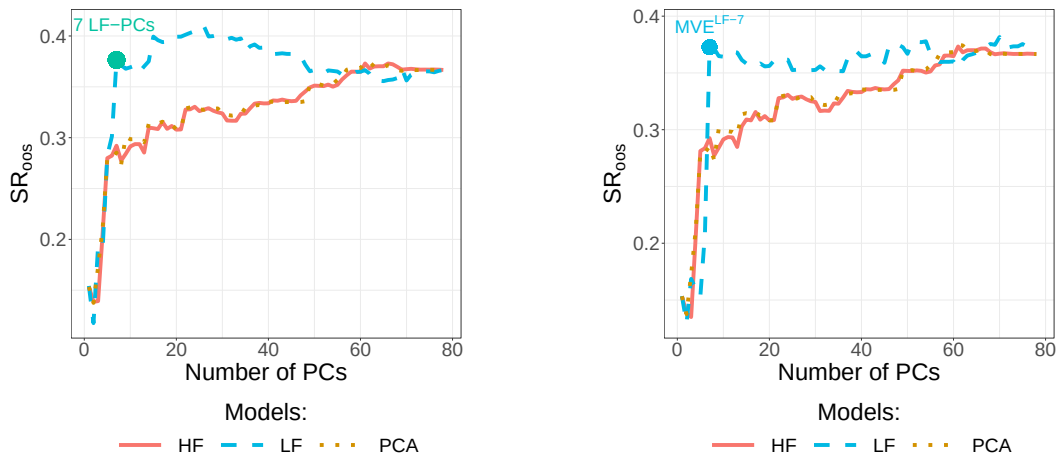
This graph zooms in the OOS Sharpe ratio of PCA, HF-PCA, and LF-PCA after imposing the CAPM. This figure shows the estimates using two prior Sharpe ratios, $SR_{prior} \in \{0.4, 0.5\}$.

IV.3 Alternative Cutoffs of the LF Interval

Earlier, I define the low-frequency component of an asset return as the part with a cycle length between 36 and 120 months. This section investigates whether the sparsity of LF factor models is particularly sensitive to alternative cutoffs of the LF intervals.

Panel (a) in Figure 13 defines the period of the LF component between 24 and 120 months. The OOS MVE portfolios of HF-PCA and PCA are identical to previous ones. The LF-MVE portfolios, however, are still more parsimonious. For example, The LF-MVE portfolio composed of the first seven LF-PCs earns a monthly Sharpe ratio of 0.35 in the out-of-sample. In addition, Figure A9 plots the heat-map for LF-PCA, whose pattern is virtually identical to that of Figure 4.

Bandi, Chaudhuri, Lo, and Tamoni (2021) decompose the CAPM β into frequency-dependent components, and they discover that only the component in the LF frequency with a period between 32 and 64 months can price conventional Fama-French portfolios. Motivated by their results, I define the period of the LF (HF) component as $\tau^{LF} \in [32, 64]$ months ($\tau^{HF} \in [2, 32)$). Panel (b) in Figure 13 shows that the monthly Sharpe ratio of a seven LF factor model is slightly higher than 0.37, whereas I still demand many HF-PCs to span the asset space of 78 test assets in the out-of-sample. In short, the sparsity of the LF-SDF is not sensitive to alternative definitions of the LF interval.



(a) $\tau^{LF} \in [24, 120]$ **(b)** $\tau^{LF} \in [32, 64]$

Figure 13: Robustness Check: OOS Sharpe ratio of 78 portfolios, $SR_{prior} = 0.4$

This graph shows additional robustness checks of the OOS Sharpe ratio of PCA, HF-PCA, and LF-PCA. I estimate risk prices under the prior Sharpe ratio 0.4. The low-frequency (LF) interval is equal to (1) panel (a): $\tau^{HF} \in [2, 24)$ and $\tau^{LF} \in [24, 120]$, and (2) panel (b): $\tau^{HF} \in [2, 32)$ and $\tau^{LF} \in [32, 64]$.

V Conclusions

I use frequency-dependent risks to dissect the factor zoo and answer fundamental questions about what is salient for cross-sectional asset pricing. As a first step, I propose a new approach to quantify frequency-dependent risks and deliver monthly proxies for short-term and long-term systematic factors. Empirically, the SDF is sparse only in the space of low-frequency systematic factors. An economic interpretation of this finding is that investors are more risk-averse to low-frequency persistent systematic factors that drive a vast majority of long-run movements of asset returns, probably because they have long investment horizons or Epstein-Zin preference that imposes a considerable risk aversion to the long-run uncertainty. Hence, the first few largest LF latent factors capture almost the entire Sharpe ratio of the true SDF.

In addition, I confirm that none of the celebrated sparse factor models, such as the Fama-French three-factor model, or HF-SDF can explain the LF-SDF. At the same time, the LF-SDF can span the HF-SDF. Therefore, I conjecture that the SDF composed of the first several low-frequency factors is the proper benchmark SDF.

Furthermore, my paper deepens our understanding of the economics behind the factor zoo. It is common to use the largest several canonical PCs to construct the SDF. This SDF, virtually identical to the HF-SDF, is almost uncorrelated over time and captures economic risks related to discount-rate news of the market excess return, intermediary factors, jump risk, and investor sentiment. However, the HF-SDF still ignores an economically important

component of the LF-SDF. This missing component commands a sizable monthly Sharpe ratio of about 0.2 and displays a persistent conditional dynamic, as the variance ratio test shows. More importantly, it reflects only business-cycle risks related strongly to consumption and GDP growth, and it can also predict consumption and GDP growth over the next quarter.

Traditional macro-finance models emphasize persistent conditional information and use them to rationalize the asset pricing puzzles. What I observe in this paper confirms that asset returns indeed contain useful conditional information related to macro variables, but they can be identified only at low frequencies. At the same time, the tail risk, behavioral finance, and intermediary asset pricing models are also essential in understanding asset returns, but they are more relevant in short horizons.

References

Alvarez, F., and U. J. Jermann (2005): "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica*, 73(6), 1977–2016.

Ang, A., J. Chen, and Y. Xing (2006): "Downside risk," *The Review of Financial Studies*, 19(4), 1191–1239.

Bai, J. (2003): "Inferential theory for factor models of large dimensions," *Econometrica*, 71(1), 135–171.

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

Baker, M., and J. Wurgler (2006): "Investor sentiment and the cross-section of stock returns," *The Journal of Finance*, 61(4), 1645–1680.

Bandi, F. M., S. E. Chaudhuri, A. W. Lo, and A. Tamoni (2021): "Spectral factor models," *Journal of Financial Economics*.

Bandi, F. M., and A. Tamoni (2017): "Business-cycle consumption risk and asset prices," *available at SSRN 2337973*.

Bansal, R., and A. Yaron (2004): "Risks for the long run: A potential resolution of asset pricing puzzles," *The Journal of Finance*, 59(4), 1481–1509.

Barillas, F., and J. Shanken (2018): "Comparing asset pricing models," *Journal of Finance*, 73(2), 715–754.

Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century," *The Quarterly Journal of Economics*, 121(3), 823–866.

Benaych-Georges, F., and R. R. Nadakuditi (2011): "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices," *Advances in Mathematics*, 227(1), 494–521.

Breeden, D. T. (1979): "An intertemporal asset pricing model with stochastic consumption and investment opportunities," *Journal of Financial Economics (JFE)*, 7(3).

Brennan, M. J., and Y. Zhang (2020): "Capital asset pricing with a stochastic horizon," *Journal of Financial and Quantitative Analysis*, 55(3), 783–827.

Bryzgalova, S., J. Huang, and C. Julliard (2021): "Bayesian solutions for the factor zoo: We just ran two quadrillion models," *available at SSRN 3481736*.

Campbell, J. Y. (1990): "A variance decomposition for stock returns," .

——— (1999): "Asset prices, consumption, and the business cycle," *Handbook of Macroeconomics*, 1, 1231–1303.

Campbell, J. Y., and J. H. Cochrane (1999): "By force of habit: A consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107(2), 205–251.

Campbell, J. Y., S. Giglio, and C. Polk (2013): "Hard times," *The Review of Asset Pricing Studies*, 3(1), 95–132.

Campbell, J. Y., and T. Vuolteenaho (2004): "Bad beta, good beta," *American Economic Review*, 94(5), 1249–1275.

Carhart, M. M. (1997): "On persistence in mutual fund performance," *The Journal of Finance*, 52(1), 57–82.

Chamberlain, G., and M. Rothschild (1983): "Arbitrage, factor structure, and mean-variance analysis on large asset markets," *Econometrica*, 51(5), 1281–1304.

Chernov, M., L. A. Lochstoer, and S. R. Lundeby (forthcoming): "Conditional dynamics and the multi-horizon risk-return trade-off," *Review of Financial Studies*.

Cochrane, J. H. (2009): *Asset Pricing: Revised Edition*. Princeton University Press.

Cochrane, J. H., and J. Saa-Requejo (2000): "Beyond arbitrage: Good-deal asset price bounds in incomplete markets," *Journal of Political Economy*, 108(1), 79–119.

Connor, G., and R. A. Korajczyk (1986): "Performance measurement with the arbitrage pricing theory: A new framework for analysis," *Journal of Financial Economics*, 15(3), 373–394.

——— (1988): "Risk and return in an equilibrium APT: Application of a new test methodology," *Journal of Financial Economics*, 21(2), 255–289.

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

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Noise trader risk in financial markets," *Journal of Political Economy*, 98(4), 703–738.

Dew-Becker, I., and S. Giglio (2016): "Asset pricing in the frequency domain: theory and empirics," *Review of Financial Studies*, 29, 2029–68.

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

——— (2015): "A five-factor asset pricing model," *Journal of Financial Economics*, 116(1), 1–22.

Gabaix, X. (2012): "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance," *The Quarterly Journal of Economics*, 127(2), 645–700.

Giannone, D., M. Lenza, and G. E. Primiceri (2021): "Economic predictions with big data: The illusion of sparsity," *ECB Working Paper*.

Giglio, S., and D. Xiu (2021): "Asset pricing with omitted factors," *Journal of Political Economy*, 129(7), 000–000.

Giglio, S., D. Xiu, and D. Zhang (2021): "Test assets and weak factors," Discussion paper, National Bureau of Economic Research.

Golubov, A., and T. Konstantinidi (2019): "Where is the risk in value? Evidence from a market-to-book decomposition," *The Journal of Finance*, 74(6), 3135–3186.

Gupta, T., and B. Kelly (2019): "Factor momentum everywhere," *The Journal of Portfolio Management*, 45(3), 13–36.

Haddad, V., S. Kozak, and S. Santosh (2020): "Factor timing," *The Review of Financial Studies*, 33(5), 1980–2018.

Handa, P., S. P. Kothari, and C. Wasley (1989): "The relation between the return interval and betas: Implications for the size effect," *Journal of Financial Economics*, 23(1), 79–100.

Hannan, E. J. (2009): *Multiple Time Series*, vol. 38. John Wiley & Sons.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption strikes back? Measuring long-run risk," *Journal of Political Economy*, 116(2), 260–302.

Hansen, L. P., and R. Jagannathan (1991): "Implications of security market data for models of dynamic economies," *Journal of Political Economy*, 99(2), 225–262.

Hansen, L. P., and J. A. Scheinkman (2009): "Long-term risk: An operator approach," *Econometrica*, 77(1), 177–234.

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

He, Z., and A. Krishnamurthy (2013): "Intermediary asset pricing," *American Economic Review*, 103(2), 732–70.

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

Huang, D., F. Jiang, J. Tu, and G. Zhou (2015): "Investor sentiment aligned: A powerful predictor of stock returns," *The Review of Financial Studies*, 28(3), 791–837.

Jegadeesh, N. (1990): "Evidence of predictable behavior of security returns," *The Journal of Finance*, 45(3), 881–898.

Kan, R., and C. Zhang (1999): "GMM tests of stochastic discount factor models with useless factors," *Journal of Financial Economics*, 54(1), 103–127.

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

Kozak, S., S. Nagel, and S. Santosh (2018): "Interpreting factor models," *The Journal of Finance*, 73(3), 1183–1223.

——— (2020): "Shrinking the cross-section," *Journal of Financial Economics*, 135(2), 271–292.

Lettau, M., and S. Ludvigson (2001a): "Consumption, aggregate wealth, and expected stock returns," *Journal of Finance*, 56(3), 815–849.

——— (2001b): "Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying," *Journal of Political Economy*, 109(6), 1238–1287.

Lettau, M., and M. Pelger (2020a): "Estimating latent asset-pricing factors," *Journal of Econometrics*, 218(1), 1–31.

——— (2020b): "Factors that fit the time series and cross-section of stock returns," *The Review of Financial Studies*, 33(5), 2274–2325.

Liew, J., and M. Vassalou (2000): "Can book-to-market, size and momentum be risk factors that predict economic growth?," *Journal of Financial Economics*, 57(2), 221–245.

Lintner, J. (1965): "Security prices, risk, and maximal gains from diversification," *The Journal of Finance*, 20(4), 587–615.

Lucas, R. E. (1978): "Asset prices in an exchange economy," *Econometrica: Journal of the Econometric Society*, pp. 1429–1445.

Markowitz, H. (1952): "Portfolio Selection in The Journal of Finance Vol. 7," .

McLean, R. D., and J. Pontiff (2016): "Does academic research destroy stock return predictability?," *The Journal of Finance*, 71(1), 5–32.

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

Neuhierl, A., and R. T. Varneskov (2021): "Frequency dependent risk," *Journal of Financial Economics*, 140(2), 644–675.

Newey, W. K., and K. D. West (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55(3), 703–708.

Onatski, A. (2012): "Asymptotics of the principal components estimator of large factor models with weakly influential factors," *Journal of Econometrics*, 168(2), 244–258.

Parker, J. A., and C. Julliard (2005): "Consumption risk and the cross section of expected returns," *Journal of Political Economy*, 113(1), 185–222.

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

Ross, S. A. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13(3), 341–360.

Rubinstein, M. (1976): "The valuation of uncertain income streams and the pricing of options," *The Bell Journal of Economics*, pp. 407–425.

Sharpe, W. F. (1964): "Capital asset prices: A theory of market equilibrium under conditions of risk," *The Journal of Finance*, 19(3), 425–442.

Shiller, R. C. (2000): "Irrational exuberance," *Philosophy and Public Policy Quarterly*, 20(1), 18–23.

Tu, J., and G. Zhou (2011): "Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies," *Journal of Financial Economics*, 99(1), 204–215.

Appendices

A Additional Details on Frequency Domain Analysis

A.1 Spectral Representation Theorem

Theorem A1 (Spectral Representation Theorem, Hannan (2009)) *Suppose x_t is a mean-zero covariance stationary process, with the spectral distribution function $F(\omega)$ such that its auto-covariance function $\Sigma_x(h)$ can be expressed as*

$$\Sigma_x(h) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp\{2\pi i\omega h\} dF(\omega),$$

where $F(\omega)$ is non-decreasing, $F(-\frac{1}{2}) = 0$ and $F(\frac{1}{2}) = \Sigma_x(0)$. Then there exists a complex-value stochastic process $z(\omega)$, $\omega \in [-\frac{1}{2}, \frac{1}{2}]$, having stationary uncorrelated increments, such that x_t can be written as the stochastic integral

$$x_t = \int_{-\frac{1}{2}}^{\frac{1}{2}} \exp\{-2\pi i\omega t\} dz(\omega),$$

where $\mathrm{Var}[z(\omega_2) - z(\omega_1)] = F(\omega_2) - F(\omega_1)$. Furthermore, the Spectral Representation Theorem can be extended to multivariate case.

Suppose that $x_t(\omega)$ satisfies the differential equation: $x_t(\omega)d\omega = \exp\{-2\pi i\omega t\}dz(\omega)$. The Spectral Representation Theorem implies that $x_t(\omega)$ is uncorrelated at different frequencies, and x_t is decomposed as an equally weighted average of $x_t(\omega)$, i.e., $x_t = \int_{-\frac{1}{2}}^{\frac{1}{2}} x_t(\omega)d\omega$. Therefore, I can represent the variance of x_t as $\mathrm{Var}(x_t) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathrm{Var}[x_t(\omega)]d\omega$, where $\mathrm{Var}[x_t(\omega)]$ is the contribution from the frequency-ω component.

Suppose that $\boldsymbol{X_t}$ is a two-dimensional time series, for example, $\boldsymbol{X_t} = (x_{1t}, x_{2t})^\top$, with the auto-covariance matrix $\boldsymbol{\Sigma_X}(h)$. According to the Spectral Representation Theorem, the cross-spectrum $f_{x_1,x_2}(\omega)$ that satisfies $dF_{12}(\omega) = f_{x_1,x_2}(\omega)d\omega$ can be interpreted as the covariance between the frequency-ω components of $x_{1,t}$ and $x_{2,t}$. Next, I will consider a linear transformation of $\boldsymbol{X_t}$. Let a and b be arbitrary real numbers, and define y_t as $y_t = ax_{1,t} + bx_{2,t}$.

The spectral density function of y_t is

$$\begin{aligned}
f_y(\omega) &= \sum_{h=-\infty}^{\infty} \text{Cov}(y_{t+h}, y_t) \exp\{-2\pi i h\omega\} \\
&= \sum_{h=-\infty}^{\infty} \text{Cov}(ax_{1,t+h} + bx_{2,t+h}, ax_{1,t} + bx_{2,t}) \exp\{-2\pi i h\omega\} \\
&= \sum_{h=-\infty}^{\infty} \left[a^2\text{Var}(x_{1,t}) + b^2\text{Var}(x_{2,t}) + ab\text{Cov}(x_{1,t+h}, x_{2,t}) + ab\text{Cov}(x_{1,t}, x_{2,t+h})\right] \exp\{-2\pi i h\omega\} \\
&= a^2 f_{x_1}(\omega) + b^2 f_{x_2}(\omega) + ab f_{x_1,x_2}(\omega) + ab f_{x_2,x_1}(\omega) \\
&= a^2 f_{x_1}(\omega) + b^2 f_{x_2}(\omega) + 2ab\mathcal{R}[f_{x_1,x_2}(\omega)],
\end{aligned}$$

where the last equality makes use of the fact that $f_{x_1,x_2}(\omega) = f_{x_2,x_1}(-\omega)$ and $f_{x_2,x_1}(\omega) + f_{x_2,x_1}(-\omega) = 2\mathcal{R}[f_{x_1,x_2}(\omega)]$. There are two implications. First, I can interpret the real part of the cross-spectrum as the covariance between the frequency-ω components of $x_{1,t}$ and $x_{2,t}$. Second, I need to focus only on the real part of the cross-spectrum. This paper aims to extract PCs at different frequencies. For example, the largest PC chooses a unitary linear transformation of $\boldsymbol{X_t}$ such that its variance is maximized.

A.2 Discrete Fourier Transform (DFT)

Given data $\boldsymbol{R_1}, \ldots, \boldsymbol{R_T}$, DFT and its inverse (**IDFT**) are defined as

$$\boldsymbol{d(\omega_j)} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{R_t} \exp\{-2\pi i \omega_j t\},\ \omega_j = \frac{j}{T},\ j = 0, 1, \ldots, T-1, \tag{27}$$

$$\boldsymbol{R_t} = \frac{1}{\sqrt{T}} \sum_{j=0}^{T-1} \boldsymbol{d(\omega_j)} \exp\{2\pi i \omega_j t\},\ t = 1, \ldots, T. \tag{28}$$

Let's define the frequency-ω_j component of asset returns: $\boldsymbol{R_t}(\omega_j) = \frac{1}{\sqrt{T}}\boldsymbol{d(\omega_j)} \exp\{2\pi i \omega_j t\}$. A distinguishing feature of the aforementioned decomposition is that two components from distinct frequencies are uncorrelated by construction; that is, $\mathbf{Cov}_T\left(\boldsymbol{R_t}(\omega_j)\boldsymbol{R_t}(\omega_k)^\top\right) = \mathbf{0_{N\times N}}$ if $j \neq k$, or $\hat{\boldsymbol{f}}_{\boldsymbol{R}}(\omega_j)$ if $j = k$. The intuition is that DFT decomposes $\boldsymbol{R_t}$ into orthogonal frequency-dependent parts.

Moreover, $\boldsymbol{d}(\omega_j)\boldsymbol{d}(\omega_j)^\star = \sum_{h=-(n-1)}^{n-1} \hat{\boldsymbol{\Sigma}}_{\boldsymbol{R}}(h) \exp\{-2\pi i \omega_j h\} = \hat{\boldsymbol{f}}_{\boldsymbol{R}}(\omega_j)$, where $\boldsymbol{d}(\omega_j)^\star$ is the conjugate transpose operation of $\boldsymbol{d}(\omega_j)$. Therefore, we can estimate the frequency density matrix of asset returns via DFT. In practice, researchers often use a fast Fourier transform (**FFT**) algorithm to compute the transformations in replace of DFT rapidly. Figure A3 is a

simple example of DFT.

B Proofs

B.1 Proof of proposition 1

Since $\boldsymbol{e_{t+1}}$ is conditionally independent, its conditional expectation is always zero: For $h > 0$, $\mathbb{E}[\boldsymbol{e_{t+h}} \mid \mathbb{I}_t] = 0$, where $\mathbb{I}_t$ denotes the conditional information at time t. If $h > 0$, the auto-covariance matrix of $\boldsymbol{e_{t+1}}$ is

$$\boldsymbol{\Sigma_e}(h) = \mathbb{E}(\boldsymbol{e_{t+h}} \boldsymbol{e_t^\top}) = \mathbb{E}[\mathbb{E}(\boldsymbol{e_{t+h}} \boldsymbol{e_t^\top} \mid \mathbb{I}_t)] = \mathbb{E}[\mathbb{E}(\boldsymbol{e_{t+h}} \mid \mathbb{I}_t) \boldsymbol{e_t^\top}] = \boldsymbol{0_{N \times N}},$$

which implies that the spectral density matrix of $\boldsymbol{e_{t+1}}$ is

$$\boldsymbol{f_e}(\omega) = \sum_{h=-\infty}^{\infty} \boldsymbol{\Sigma_e}(h) \exp\{-2\pi i h \omega\} = \boldsymbol{\Sigma_e}(0) = \boldsymbol{\Sigma_e}.$$

Therefore, even though $\boldsymbol{e_{t+1}}$ and $\boldsymbol{f_{t+1}}$ can follow stochastic volatility processes, their spectral density matrices are constant across frequencies.

In addition, $\boldsymbol{F_t}$ and $\boldsymbol{e_t}$ are orthogonal, so I can represent the covariance matrix of $\boldsymbol{R_t}$ as $\boldsymbol{\Sigma_R} = \boldsymbol{\beta} \boldsymbol{\Sigma_F} \boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e}$. Suppose that $\boldsymbol{f_F}(\omega)$ is the spectral density matrix of $\boldsymbol{F_t}$: $\boldsymbol{\Sigma_F} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{f_F}(\omega) d\omega$. This implies the following spectral decomposition of $\boldsymbol{\Sigma_R}$:

$$\boldsymbol{\Sigma_R} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \boldsymbol{\beta} \boldsymbol{f_F}(\omega) \boldsymbol{\beta}^\top d\omega + \boldsymbol{\Sigma_e} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[\boldsymbol{\beta} \boldsymbol{f_F}(\omega) \boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e}\right] d\omega.$$

Due to the uniqueness of the spectral density matrix, the spectral density matrix of $\boldsymbol{R_t}$ is $\boldsymbol{f_R}(\omega) = \boldsymbol{\beta} \boldsymbol{f_F}(\omega) \boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e}$. Similarly, I show that $\boldsymbol{f_F}(\omega) = \boldsymbol{\Sigma_f} + \boldsymbol{\Phi_X} \boldsymbol{f_X}(\omega) \boldsymbol{\Phi_X^\top}$. Therefore, I rewrite the spectral density matrix of $\boldsymbol{R_t}$ as follows:

$$\boldsymbol{f_R}(\omega) = \boldsymbol{\beta} \boldsymbol{\Sigma_f} \boldsymbol{\beta}^\top + \boldsymbol{\Sigma_e} + \boldsymbol{\beta_X} \boldsymbol{f_X}(\omega) \boldsymbol{\beta_X^\top}, \quad \boldsymbol{\beta_X} = \boldsymbol{\beta} \boldsymbol{\Phi_X}.$$

B.2 Proof of proposition 2

I can derive the unconditional variance of the linear SDF as follows:

$$\begin{aligned}
\text{Var}(\mathcal{M}_{t+1}) &= \boldsymbol{b}^\top \text{Var}(\boldsymbol{f_{t+1}})\boldsymbol{b} + \boldsymbol{b}^\top \boldsymbol{\Phi_X} \text{Var}(\boldsymbol{X_t}) \boldsymbol{\Phi_X}^\top \boldsymbol{b} \\
&= \boldsymbol{b}^\top \text{Var}(\boldsymbol{f_{t+1}})\boldsymbol{b} + \boldsymbol{b_X}^\top \text{Var}(\boldsymbol{X_t}) \boldsymbol{b_X} \\
&= \boldsymbol{b}^\top \text{Var}(\boldsymbol{f_{t+1}})\boldsymbol{b} + \sum_{j=1}^{p} b_{X,i}^2 \text{Var}(X_{jt}) \\
&= \boldsymbol{b}^\top \text{Var}(\boldsymbol{f_{t+1}})\boldsymbol{b} + \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{j=1}^{p} b_{X,i}^2 f_{X_i}(\omega) d\omega,
\end{aligned}$$

where the third equality uses the fact that state variables are assumed to be uncorrelated, and the last step uses the spectral decomposition of each state variable X_i. Since the spectral density function is unique, $f_{\mathcal{M}}(\omega) = \boldsymbol{b}^\top \text{Var}(\boldsymbol{f_{t+1}})\boldsymbol{b} + \sum_{j=1}^{p} b_{X,i}^2 f_{X_i}(\omega)$.

B.3 Derivation of objective function in equation (16)

This section derives the objective function in equation (16) under a more general distributional assumption for pricing errors and risk prices. I consider only the cross-sectional regression, conditional on the observed expectation and covariance of $\boldsymbol{F_t}$ as follows:

$$\boldsymbol{\mu_F} = \boldsymbol{\Sigma_F} \boldsymbol{b} + \boldsymbol{\alpha}, \quad \boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{0_N}, \sigma^2 \boldsymbol{\Sigma_F}).$$

Therefore, the only unknowns are $\boldsymbol{b}$ and σ^2. Pástor and Stambaugh (2000) and Barillas and Shanken (2018) also make a similar distributional assumption for $\boldsymbol{\alpha}$. Intuitively, σ^2 reflects investors' uncertainty about mispricing: When σ^2 is close to zero, the asset pricing model is almost correct. In contrast, if σ^2 is infinity, the factor model is useless, as it entirely fails to explain risk premia.

Furthermore, I assign a normal prior for risk prices: $\boldsymbol{b} \sim \mathcal{N}(\boldsymbol{0_K}, \frac{\psi\sigma^2}{\tau} \boldsymbol{I_K})$, $\tau = \text{Tr}\left[\boldsymbol{\Sigma_F}\right]$, and $\boldsymbol{b}$ is uncorrelated with $\boldsymbol{\alpha}$. Under such a prior distribution, the prior expectation on the squared Sharpe ratio of factor returns implied by the asset pricing model is equal to

$$\mathbb{E}_{prior}[SR_F^2] = \mathbb{E}_{prior}[\boldsymbol{b}^\top \boldsymbol{\Sigma_F} \boldsymbol{b}] = \sum_{k=1}^{K} \sigma_{F,k}^2 \mathbb{E}_{prior}[b_k^2] = \frac{\psi\sigma^2}{\tau} \text{Tr}\left[\boldsymbol{\Sigma_F}\right] = \psi\sigma^2.$$

Next, I decompose the expected squared Sharpe ratio of factor returns as follows:

$$\begin{aligned}\mathbb{E}_{prior}[\boldsymbol{\mu}_{\boldsymbol{F}}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{\mu}_{\boldsymbol{F}}] &= \mathbb{E}_{prior}[(\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b}+\boldsymbol{\alpha})^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}(\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b}+\boldsymbol{\alpha})] \\ &= \mathbb{E}_{prior}[\boldsymbol{b}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b}] + \mathbb{E}_{prior}[\boldsymbol{\alpha}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{\alpha}] \\ &= \psi\sigma^2 + N\sigma^2 = (\psi+N)\sigma^2;\end{aligned}$$

therefore, $\mathbb{E}_{prior}[\boldsymbol{\mu}_{\boldsymbol{F}}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{\mu}_{\boldsymbol{F}}]$ is the sum of $\mathbb{E}_{prior}[\boldsymbol{b}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b}]$ and $\mathbb{E}_{prior}[\boldsymbol{\alpha}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{\alpha}]$, where the former is the contribution from the SDF. Also, I derive the expected squared Sharpe ratio of the SDF as follows:

$$\mathbb{E}_{prior}[SR_F^2] = \frac{\psi}{\psi+N}\mathbb{E}_{prior}[\boldsymbol{\mu}_{\boldsymbol{F}}^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}\boldsymbol{\mu}_{\boldsymbol{F}}],$$

so a larger ψ implies higher prior Sharpe ratio of the SDF. Under the above assumptions, the posterior distribution of $\boldsymbol{b}$, conditional on $(\boldsymbol{\mu}_{\boldsymbol{F}}, \boldsymbol{\Sigma}_{\boldsymbol{F}})$, is

$$\begin{aligned}p(\boldsymbol{b} \mid \boldsymbol{\mu}_{\boldsymbol{F}}, \boldsymbol{\Sigma}_{\boldsymbol{F}}) &\propto \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{\mu}_{\boldsymbol{F}}-\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b})^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}(\boldsymbol{\mu}_{\boldsymbol{F}}-\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b})\right\}\exp\left\{-\frac{\tau}{2\psi\sigma^2}\boldsymbol{b}^{\top}\boldsymbol{b}\right\} \\ &\propto \exp\left\{-\frac{1}{2\sigma^2}\left[(\boldsymbol{\mu}_{\boldsymbol{F}}-\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b})^{\top}\boldsymbol{\Sigma}_{\boldsymbol{F}}^{-1}(\boldsymbol{\mu}_{\boldsymbol{F}}-\boldsymbol{\Sigma}_{\boldsymbol{F}}\boldsymbol{b}) + \frac{\tau}{\psi}\boldsymbol{b}^{\top}\boldsymbol{b}\right]\right\}.\end{aligned}$$

Now let $v_2 = \frac{\tau}{\psi} = \frac{\tau\sigma^2}{\mathbb{E}_{prior}[SR_F^2]}$. Therefore, the posterior mode of $\boldsymbol{b}$ is the solution to the objective function in equation (16).

To compare with Kozak, Nagel, and Santosh (2020), this paper adopts a similar strategy, which assumes $\sigma^2 = \frac{1}{T}$, so $v_2 = \frac{\tau}{\psi} = \frac{\tau}{T\times\mathbb{E}_{prior}[SR_F^2]}$. Last but not least, the assumption of $\sigma^2 = \frac{1}{T}$ changes only the prior Sharpe ratio implied by the SDF. In this paper, I show the empirical results across a wide range of prior Sharpe ratios. More importantly, empirical results are robust when I estimate the model with reasonable prior *monthly* Sharpe ratios, for example, between 0.3 and 0.6.

C Additional Tables

Table A1: Definition of Variables

Variable	Definition	Data Source
R_{mt}	Monthly market excess return	CRSP database
TY_t	Monthly term yield spread, the difference between the log yield on the 10-year U.S. Constant Maturity Bond and the log yield on the three-month U.S. treasury bills	FRED
PE_t	Monthly pricing-earnings ratio (PE) from Shiller (2000)	Robert Shiller's website
VS_t	Small-stock value spread that is the difference in $\log(\frac{BE}{ME})$ between the small high-book-to-market portfolio and the small low-book-to-market portfolio	Ken French's website
DEF_t	Monthly default spread, the difference between the log yield on Moody's BAA and AAA bonds	FRED
C_t^{nd}	Quarterly real nondurable consumption growth per capita	Table 7.1 in BEA
GDP_t	Quarterly real GDP growth per capita	Table 7.1 in BEA
N_t^{CF}	Monthly cash-flow news in Campbell and Vuolteenaho (2004) Four state variables in VAR(1): $(\log(R_{mt}), TY_t, PE_t, VS_t)$	Estimated by this paper
N_t^{DR}	Monthly discount-rate news in Campbell and Vuolteenaho (2004) Four state variables in VAR(1): $(\log(R_{mt}), TY_t, PE_t, VS_t)$	Estimated by this paper
$N_t^{CF,2}$	Monthly cash-flow news in Campbell, Giglio, and Polk (2013) Five state variables in VAR(1): $(\log(R_{mt}), TY_t, PE_t, VS_t, DEF_t)$	Estimated by this paper
$N_t^{DR,2}$	Monthly discount-rate news in Campbell, Giglio, and Polk (2013) Five state variables in VAR(1): $(\log(R_{mt}), TY_t, PE_t, VS_t, DEF_t)$	Estimated by this paper
HKM_t^I	Monthly nontradable intermediary factor in He, Kelly, and Manela (2017)	Author's website
HKM_t^{II}	Monthly tradable intermediary factor in He, Kelly, and Manela (2017)	Author's website
HKM_{qt}^I	Quarterly nontradable intermediary factor in He, Kelly, and Manela (2017)	Author's website
HKM_{qt}^{II}	Quarterly tradable intermediary factor in He, Kelly, and Manela (2017)	Author's website
VXO_t	the VXO index	WRDS database
BW_t	Sentiment index in Baker and Wurgler (2006)	Dashan Huang's website
HJTZ_t	Sentiment index in Huang, Jiang, Tu, and Zhou (2015)	Dashan Huang's website
VXO_t^{ar1}	AR(1) shock in VXO_t: $VXO_t - \rho \times VXO_{t-1}$	Estimated by this paper
BW_t^{ar1}	AR(1) shock in BW_t: $BW_t - \rho \times BW_{t-1}$	Estimated by this paper
HJTZ_t^{ar1}	AR(1) shock in HJTZ_t: $\text{HJTZ}_t - \rho \times \text{HJTZ}_{t-1}$	Estimated by this paper

Table A2: 39 Firm Characteristics in Kozak, Nagel, and Santosh (2020)

Category	Characteristics
Reversal	lrrev, strev, indmomrev, indrrev, indrrevlv
Momentum	mom, mom12, indmom, momrev
Value	value, valuem, divp, ep, cfp, sp
Investment	inv, invcap, igrowth, growth, noa
Profitability	prof, roaa, roea, gmargins
Value interaction	valmom, valmomprof, valprof
Trading frictions	ivol, shvol, aturnover
Others	size, price, accruals, ciss, lev, season, sgrowth, nissa, dur

Table A3: Do celebrated models explain HF and LF risks? ($SR_{prior} = 0.5$)

		Panel (A). $\mathbf{MVE}_t^{HF}$				Panel (B). $\mathbf{MVE}_t^{LF}$			
		7 PCs	8 PCs	9 PCs	10 PCs	7 PCs	8 PCs	9 PCs	10 PCs
CAPM	α	1.08%	1.03%	1.15%	1.18%	1.47%	1.46%	1.45%	1.47%
	t-stat I	(2.98)	(2.49)	(2.91)	(3.03)	(3.66)	(3.84)	(3.83)	(3.88)
	t-stat II	(3.15)	(2.75)	(3.20)	(3.33)	(4.69)	(4.84)	(4.84)	(4.83)
	R^2_{adj}	3.89%	2.29%	0.13%	0.07%	2.79%	2.05%	1.63%	0.83%
FF3	α	0.85%	0.80%	0.95%	0.99%	1.37%	1.37%	1.36%	1.38%
	t-stat I	(4.87)	(3.79)	(4.12)	(4.27)	(4.31)	(4.52)	(4.51)	(4.62)
	t-stat II	(4.79)	(4.05)	(4.44)	(4.58)	(5.75)	(5.90)	(5.92)	(6.00)
	R^2_{adj}	36.19%	33.64%	29.11%	23.11%	8.85%	7.56%	6.53%	6.19%
Carhart4	α	0.47%	0.41%	0.59%	0.61%	0.92%	0.91%	0.90%	0.92%
	t-stat I	(3.28)	(2.57)	(3.01)	(3.10)	(3.66)	(3.80)	(3.78)	(3.83)
	t-stat II	(2.89)	(2.37)	(3.02)	(3.09)	(4.43)	(4.50)	(4.49)	(4.48)
	R^2_{adj}	62.35%	62.3%	53.52%	50.1%	46.84%	45.86%	46.91%	45.52%
FF5	α	0.47%	0.45%	0.49%	0.55%	0.95%	0.92%	0.90%	0.90%
	t-stat I	(2.62)	(2.37)	(2.80)	(3.04)	(3.12)	(3.15)	(3.09)	(3.10)
	t-stat II	(2.53)	(2.48)	(3.23)	(3.42)	(4.02)	(3.98)	(3.90)	(3.95)
	R^2_{adj}	45.98%	44.58%	46.93%	39.31%	21.81%	22.4%	22.17%	22.69%
Q4	α	0.35%	0.24%	0.27%	0.34%	0.84%	0.81%	0.78%	0.79%
	t-stat I	(2.17)	(1.29)	(1.42)	(1.75)	(2.73)	(2.79)	(2.71)	(2.77)
	t-stat II	(1.69)	(1.11)	(1.29)	(1.61)	(3.34)	(3.33)	(3.24)	(3.33)
	R^2_{adj}	39.04%	36.74%	44.05%	39.22%	25.73%	26.12%	26.91%	27.12%

This table tests whether five sparse factor models proposed in past literature can explain the MVE portfolios composed of latent factors. I construct the MVE portfolios using the first seven to 10 latent factors following the same steps as in the section III.4. I estimate the factors' risk prices under the prior Sharpe ratio of 0.5. The five benchmark models include (1) CAPM, (2) Fama and French (1993) three factors (FF3), (3) Fama and French (2015) five factors (FF5), (4) Carhart (1997) four factors (Carhart4), and (5) Hou, Xue, and Zhang (2015) four factors (Q4). I report three test-statistic in table 2: (1) α, (2) t-statistics of α, and (3) adjusted R-squared, denoted as R^2_{adj}. To control for the serial dependence of pricing errors, I use Newey and West (1987) standard errors with both 36 lags (t-stat I) and 12 lags (t-stat II).

Table A4: Correlation among $\mathcal{M}_t^{LF}$, $\mathcal{M}_t^{HF}$, $\mathcal{M}_t^{missing}$, and $\mathcal{M}_t^{unpriced}$, $SR_{prior} = 0.5$

Corr.	$\mathcal{M}_t^{LF}$	$\mathcal{M}_t^{unpriced}$	$\mathcal{M}_t^{HF}$	$\mathcal{M}_t^{missing}$	SR	t-stat (optimal lags)
$\mathcal{M}_t^{LF}$	1.00				0.378	6.24
$\mathcal{M}_t^{unpriced}$	0.00	1.00			0.014	0.25
$\mathcal{M}_t^{HF}$	0.79	0.61	1.00		0.290	4.65
$\mathcal{M}_t^{missing}$	0.61	-0.79	0.00	1.00	0.244	4.62

This table tests whether the LF-MVE portfolio can explain the HF-MVE or whether the opposite is valid. I construct the MVE portfolios using the first 7 – 10 latent factors following the same steps as in Section III.4. I estimate the factors' risk prices under the prior Sharpe ratio of 0.5. I report three test-statistic in Table 2: (1) α, (2) t-statistics of α, and (3) adjusted R-squared, denoted as R^2_{adj}. To control for the serial dependence of pricing errors, I use Newey and West (1987) standard errors with both 36 lags (t-stat I) and 12 lags (t-stat II).

Table A5: Economic Properties of HF- vs. LF-SDFs **II**

Y_t :	$N_t^{CF,2}$	$N_t^{DR,2}$	HKM_{qt}^{ntr}	HKM_{qt}^{tr}	VXO_t	BW_t	HJTZ_t	HJTZ_t^{ar1}
Panel (A): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{HF} + \beta_2 \mathcal{M}_t^{missing} + \epsilon_t$								
$\mathcal{M}_t^{HF}$	-0.134	-0.278	-0.301	-0.344	0.215	-0.193	-0.162	-0.125
t-stat I	(-1.125)	(-2.931)	(-2.642)	(-2.843)	(1.309)	(-1.641)	(-1.030)	(-1.658)
t-stat II	(-0.952)	(-2.426)	(-2.466)	(-2.695)	(1.383)	(-1.523)	(-0.912)	(-1.695)
$\mathcal{M}_t^{missing}$	-0.099	-0.048	0.206	0.193	-0.053	-0.085	-0.076	0.054
t-stat I	(-1.427)	(-0.695)	(1.480)	(1.398)	(-0.724)	(-1.743)	(-1.272)	(0.665)
t-stat II	(-1.325)	(-0.696)	(1.562)	(1.572)	(-0.669)	(-1.791)	(-1.164)	(0.787)
Panel (B): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{LF} + \beta_2 \mathcal{M}_t^{unpriced} + \epsilon_t$								
$\mathcal{M}_t^{LF}$	-0.167	-0.257	-0.146	-0.189	0.148	-0.209	-0.177	-0.074
t-stat I	(-1.560)	(-2.861)	(-1.522)	(-1.726)	(0.939)	(-1.861)	(-1.116)	(-0.965)
t-stat II	(-1.570)	(-2.929)	(-2.177)	(-2.547)	(0.947)	(-1.740)	(-0.991)	(-1.000)
$\mathcal{M}_t^{unpriced}$	0.007	-0.115	-0.334	-0.346	0.164	-0.037	-0.027	-0.114
t-stat I	(0.062)	(-1.244)	(-2.258)	(-2.271)	(1.956)	(-0.557)	(-0.426)	(-1.393)
t-stat II	(0.059)	(-1.144)	(-2.006)	(-2.153)	(2.125)	(-0.578)	(-0.427)	(-1.671)
ρ	-0.173	-0.114	-0.023	0.024	0.888	0.951	0.985	0.408
R^2_{adj}	2.78%	7.96%	13.30%	15.54%	4.90%	4.50%	3.21%	1.85%
Sample size	338	338	112	112	338	326	326	326

This table reports the results of regressing economic variables on different components of SDFs. It differs from Table 5 in following aspects: (1) I estimate cash-flow and discount-rate news including five state variables into VAR(1) regression, as in Campbell, Giglio, and Polk (2013); (2) I use quarterly intermediary factors rather than monthly ones; (3) I use the original time-series of VXO index and Baker and Wurgler (2006) sentiment index, rather than AR(1) shocks in these variables. In addition, I consider the sentiment index in Huang, Jiang, Tu, and Zhou (2015) in the last two columns.

Specifically, the dependent variables include (1) cash-flow news, (2) discount-rate news, (3) the quarterly nontradable intermediary factor, (4) the quarterly tradable intermediary factor, (5) the VXO index, (6) the Baker and Wurgler (2006) sentiment index, (7) the Huang, Jiang, Tu, and Zhou (2015) sentiment index, and (8) the AR(1) shock in the Huang, Jiang, Tu, and Zhou (2015) sentiment.

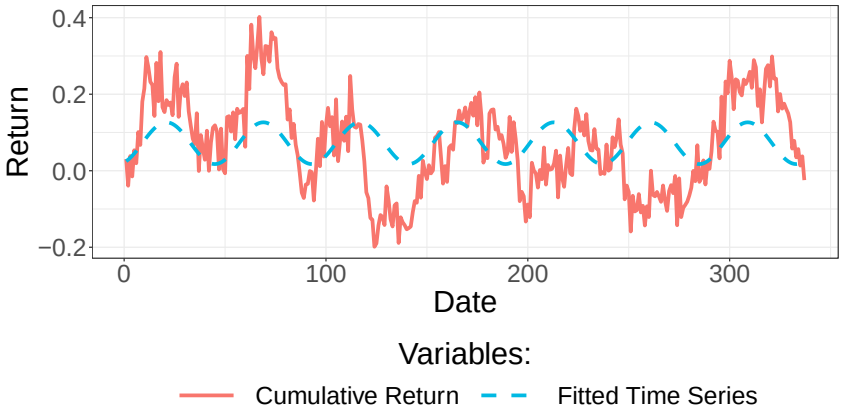
The SDFs are composed of the first seven principal components of asset returns and their risk prices are estimated under the prior Sharpe ratio equal to 0.4. I standardize both dependent and independent variables so that readers can interpret all coefficient estimates as correlations. I report two t-statistics using Newey and West (1987) standard errors with (1) 36 lags (t-stat I) and (2) 12 lags (t-stat II). In addition, I also report dependent variables' first-order autocorrelation coefficients (ρ). The monthly (quarterly) out-of-sample is from November 1991 to December 2019 (Q1 1992 – Q4 2019).

Table A6: Economic Fundamentals related to HF- vs. LF-SDFs, $SR_{prior} = 0.5$

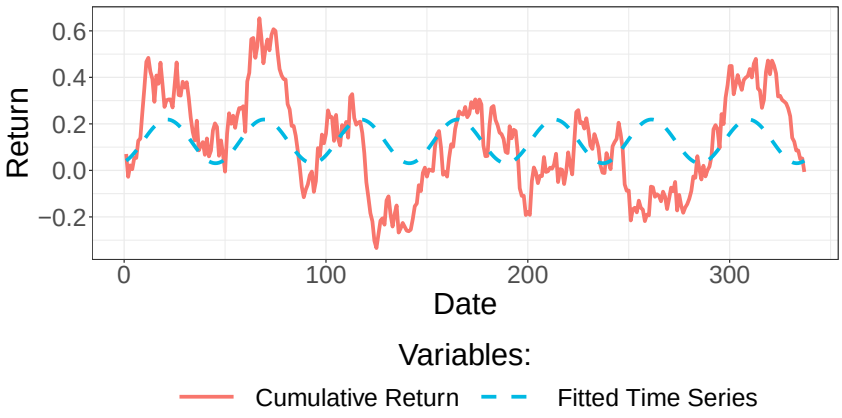
Y_t :	C_t^{nd}	C_{t+1}^{nd}	GDP_t	GDP_{t+1}	N_t^{CF}	N_t^{DR}	HKM_t^{ntr}	HKM_t^{tr}	VXO_t^{ar1}	BW_t^{ar1}
Panel (A): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{HF} + \beta_2 \mathcal{M}_t^{missing} + \epsilon_t$										
$\mathcal{M}_t^{HF}$	-0.011	0.169	-0.147	0.062	-0.062	-0.227	-0.173	-0.221	0.182	-0.152
t-stat I	(-0.113)	(1.707)	(-0.650)	(0.635)	(-0.454)	(-2.299)	(-1.879)	(-1.931)	(2.154)	(-2.618)
t-stat II	(-0.128)	(1.707)	(-0.693)	(0.694)	(-0.412)	(-2.004)	(-1.709)	(-1.863)	(2.128)	(-2.772)
$\mathcal{M}_t^{missing}$	-0.192	-0.209	-0.158	-0.192	-0.081	-0.004	0.162	0.200	-0.080	-0.014
t-stat I	(-1.614)	(-3.304)	(-1.511)	(-1.985)	(-1.057)	(-0.054)	(1.769)	(1.893)	(-1.629)	(-0.188)
t-stat II	(-1.864)	(-3.304)	(-1.594)	(-2.209)	(-1.056)	(-0.051)	(1.721)	(1.857)	(-1.635)	(-0.188)
Panel (B): $Y_t = \beta_0 + \beta_1 \mathcal{M}_t^{LF} + \beta_2 \mathcal{M}_t^{unpriced} + \epsilon_t$										
$\mathcal{M}_t^{LF}$	-0.123	0.013	-0.212	-0.064	-0.099	-0.181	-0.037	-0.052	0.094	-0.129
t-stat I	(-1.009)	(0.147)	(-0.925)	(-0.613)	(-0.895)	(-2.077)	(-0.429)	(-0.538)	(1.351)	(-1.876)
t-stat II	(-0.995)	(0.149)	(-0.965)	(-0.664)	(-0.896)	(-2.207)	(-0.475)	(-0.567)	(1.409)	(-1.871)
$\mathcal{M}_t^{unpriced}$	0.149	0.269	0.040	0.192	0.025	-0.137	-0.234	-0.294	0.175	-0.082
t-stat I	(3.217)	(3.264)	(0.415)	(2.138)	(0.201)	(-1.443)	(-2.252)	(-2.332)	(2.592)	(-1.296)
t-stat II	(2.856)	(3.281)	(0.497)	(2.418)	(0.195)	(-1.259)	(-2.050)	(-2.228)	(2.421)	(-1.304)
ρ	0.153	0.153	0.352	0.352	-0.189	-0.108	0.061	0.104	0.116	0.105
R^2_{adj}	4.91%	7.43%	6.30%	5.11%	2.76%	9.13%	7.51%	11.38%	5.80%	2.19%
Sample size	112	111	112	111	338	338	338	338	338	326

This table differs from Table 5 only in the prior Sharpe ratio that I use to estimate risk prices of latent factors. Specifically, this table sets SR_{prior} to be 0.5. See the footnote in Table 5 for details.

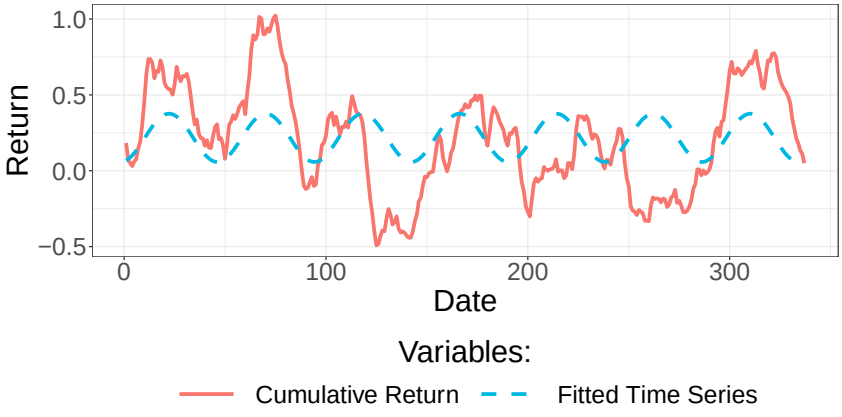
D Additional Figures



(a) Fast-Moving: AR(1) coefficient = -0.5



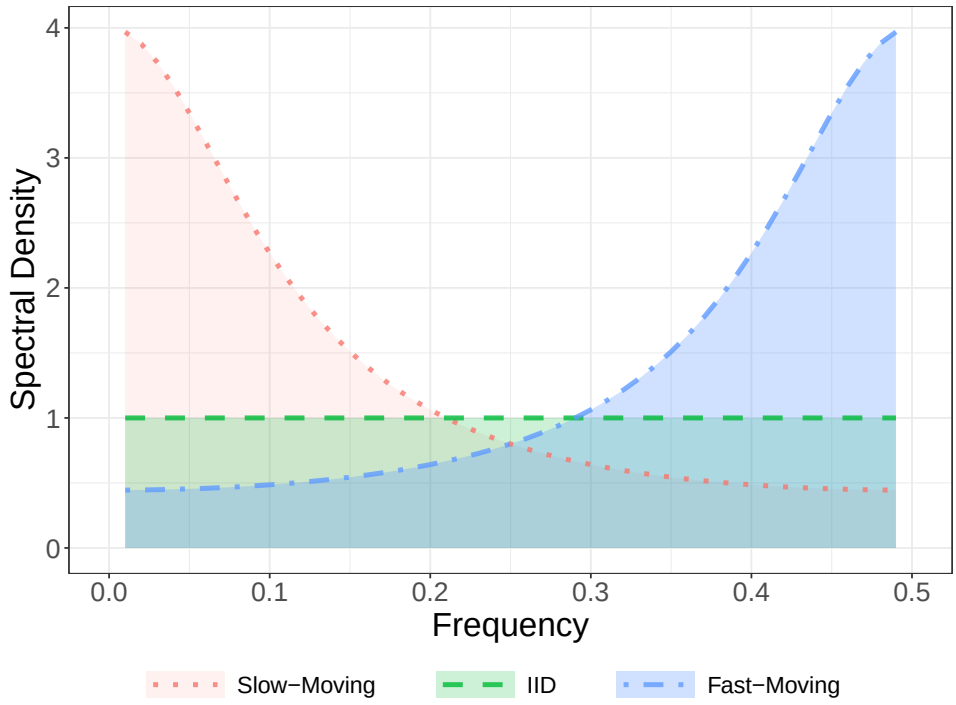
(b) IID: AR(1) coefficient = 0



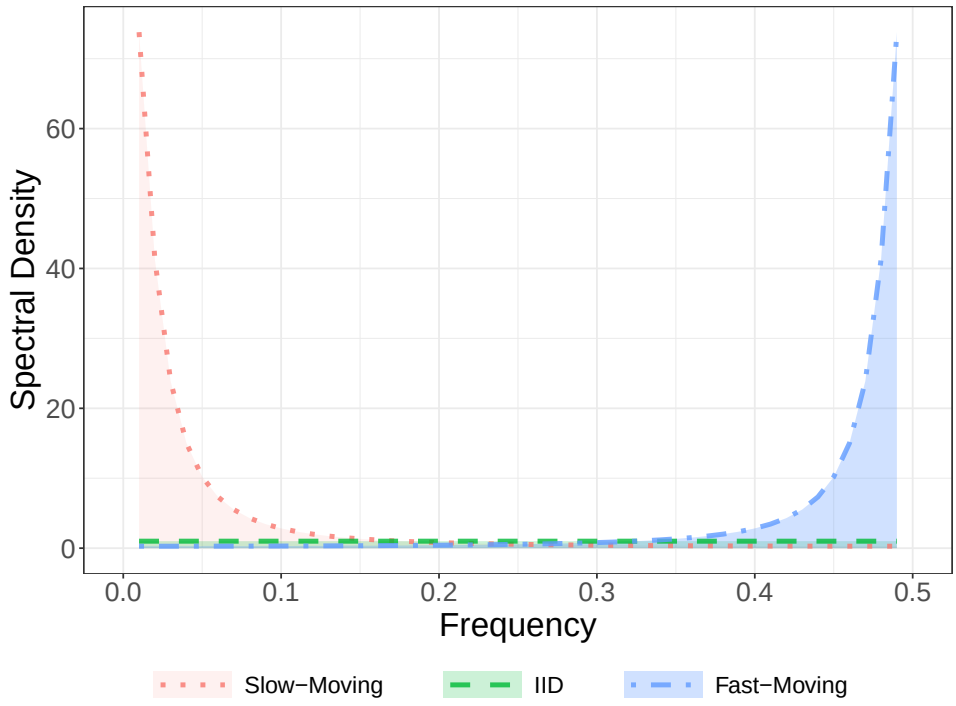
(c) Slow-Moving: AR(1) coefficient = 0.5

Figure A1: Cumulative returns in a 24-month rolling window

This graph plots the cumulative returns in a 24-month rolling window. I consider three AR(1) processes for monthly (demeaned) asset returns: $x_{t+1} = \rho_x x_t + \sqrt{1-\rho_x^2}\sigma_x \eta_{x,t+1}$, where $\sigma_x^2 = 4.5\%$, $\rho_x \in \{-0.5, 0, 0.5\}$.



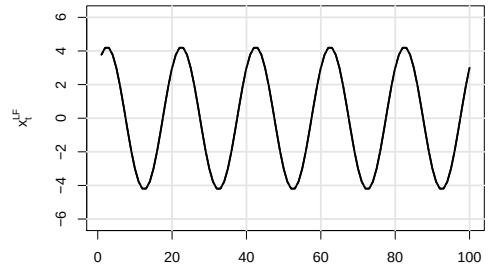
(a) $\rho_x \in \{-0.5,\ 0,\ 0.5\}$.



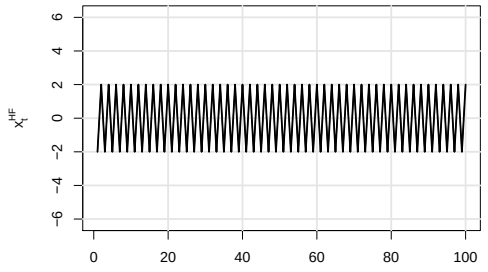
(b) $\rho_x \in \{-0.9,\ 0,\ 0.9\}$.

Figure A2: Spectral density function of AR(1) processes

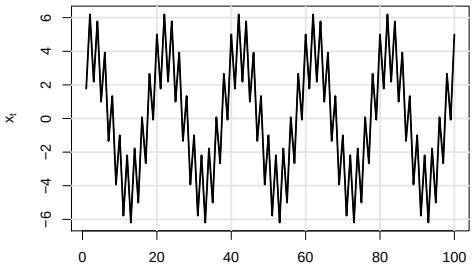
This graph plots the spectral density functions of three AR(1) processes: $x_{t+1} = \rho_x x_t + \sqrt{1-\rho_x^2}\eta_{x,t+1}$, where $\eta_{x,t+1} \overset{\text{iid}}{\sim} WN(0,1)$. When ρ_x is positive (negative), this time series is slow-moving (fast-moving).



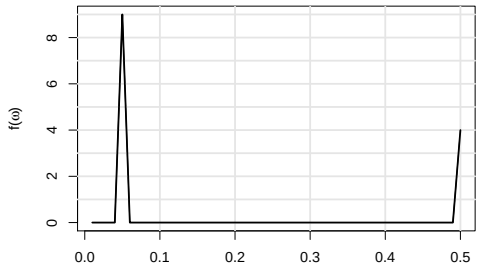
(a) $x_t^{slow} = 3\cos(2\pi t\frac{5}{100}) + 3\sin(2\pi t\frac{5}{100})$



(b) $x_t^{fast} = 2\cos(2\pi t\frac{50}{100}) + 2\sin(2\pi t\frac{50}{100})$

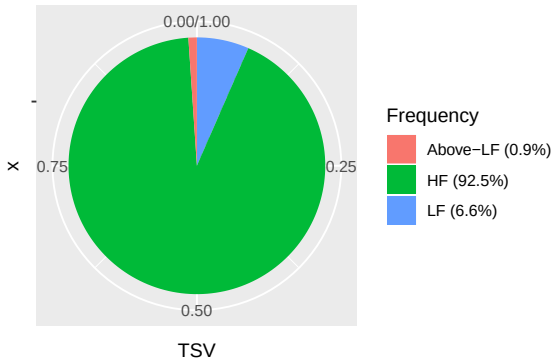


(c) $x_t = x_t^{slow} + x_t^{fast}$

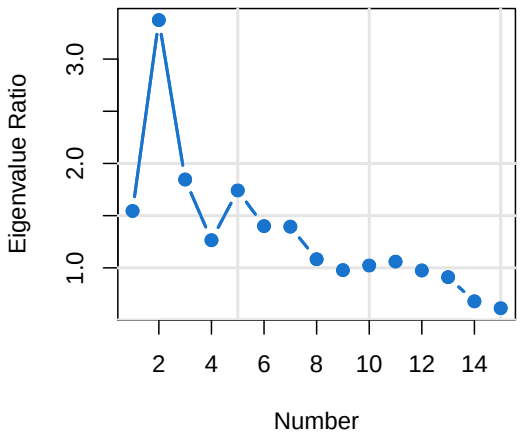


(d) Spectral density function $f_x(\omega)$, DFT

Figure A3: Example: decompose a deterministic time series via DFT



(a) TSVs explained by different components



(b) LF-eigenvalues / HF-eigenvalues

Figure A4: Time-series variations in 78 assets, subsample 2

Panel (a) plots the fraction of time-series variations in 78 asset returns explained by the HF, LF, and above-LF components. Panel (b) plots the ratio of the first 15 LF-eigenvalues over HF-eigenvalues. The sample starts from November 1991 to December 2019.

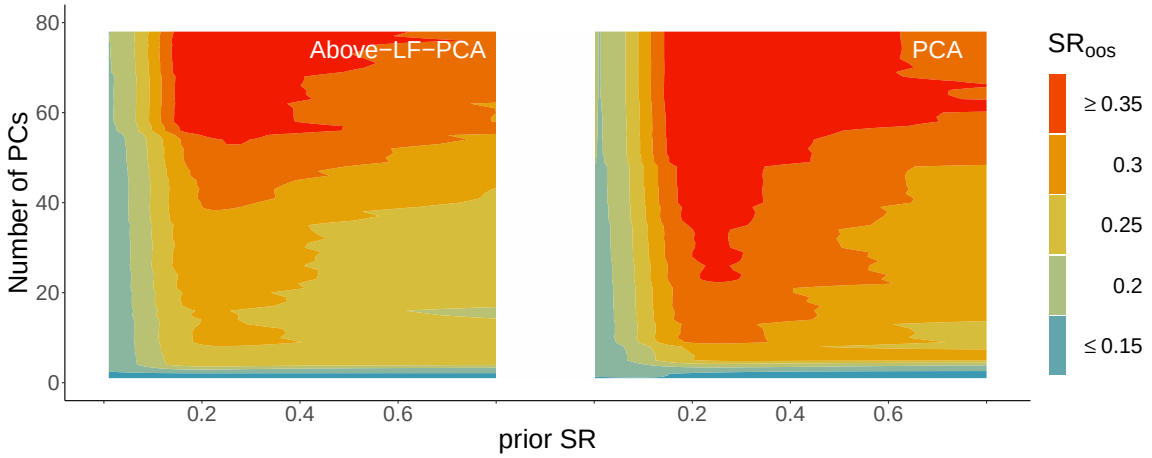


Figure A5: OOS Sharpe ratio of Above-LF-PCA and PCA, 78 test assets

This graph plots the heat-maps of the OOS Sharpe ratio of Above-LF-PCA and PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. I include the PCs into the SDF based on their ability to explain time-series variations.

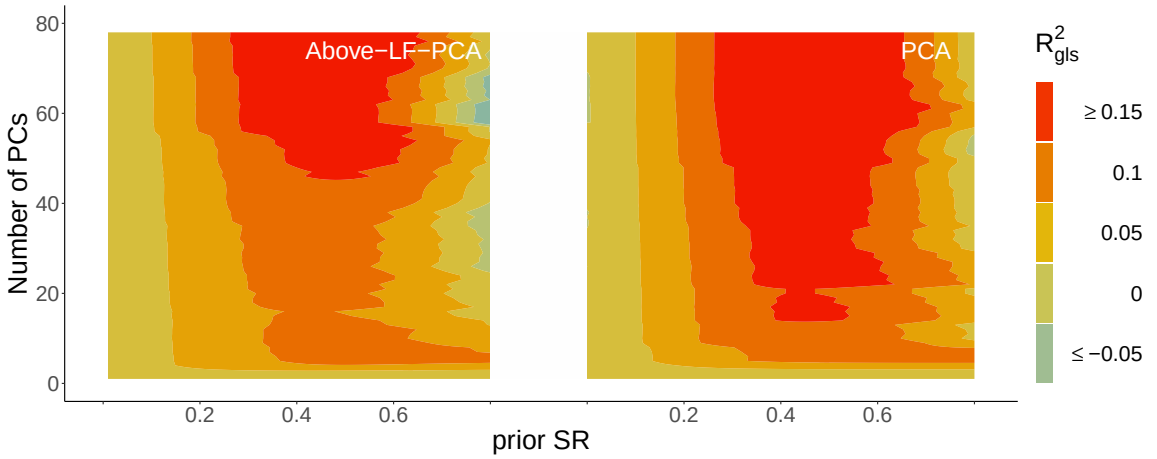


Figure A6: OOS R^2_{gls} of Above-LF-PCA and PCA, 78 test assets

This graph plots the heat-maps of the OOS R^2_{gls} of Above-LF-PCA and PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS R^2_{gls}. I include the PCs into the SDF based on their ability to explain time-series variations.

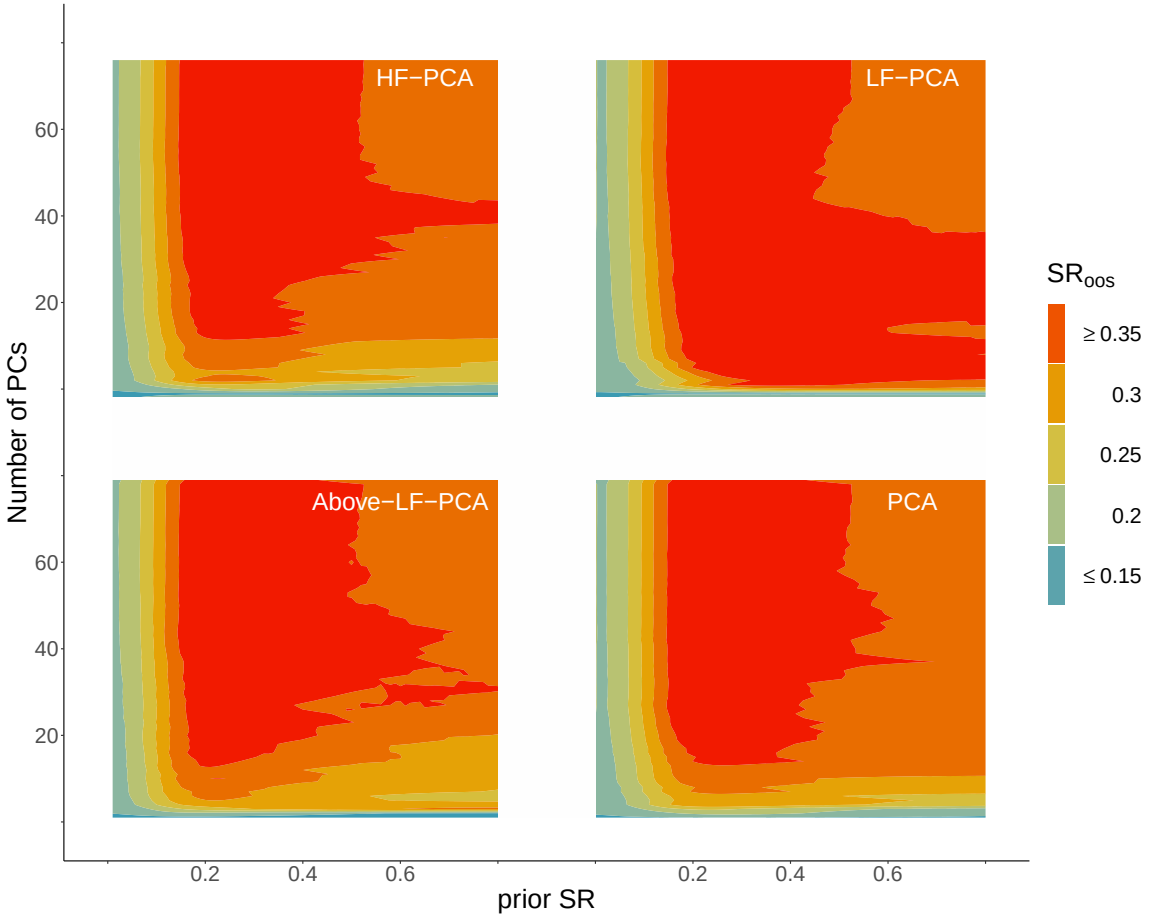


Figure A7: OOS Sharpe ratio using Kozak, Nagel, and Santosh (2020), 78 test assets

This graph plots the heat-maps of OOS Sharpe ratios of HF-PCA, LF-PCA, Above-LF-PCA and PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. The risk prices and the number of PCs entering the SDFs are determined by the Kozak, Nagel, and Santosh (2020) objective function in equation (17).

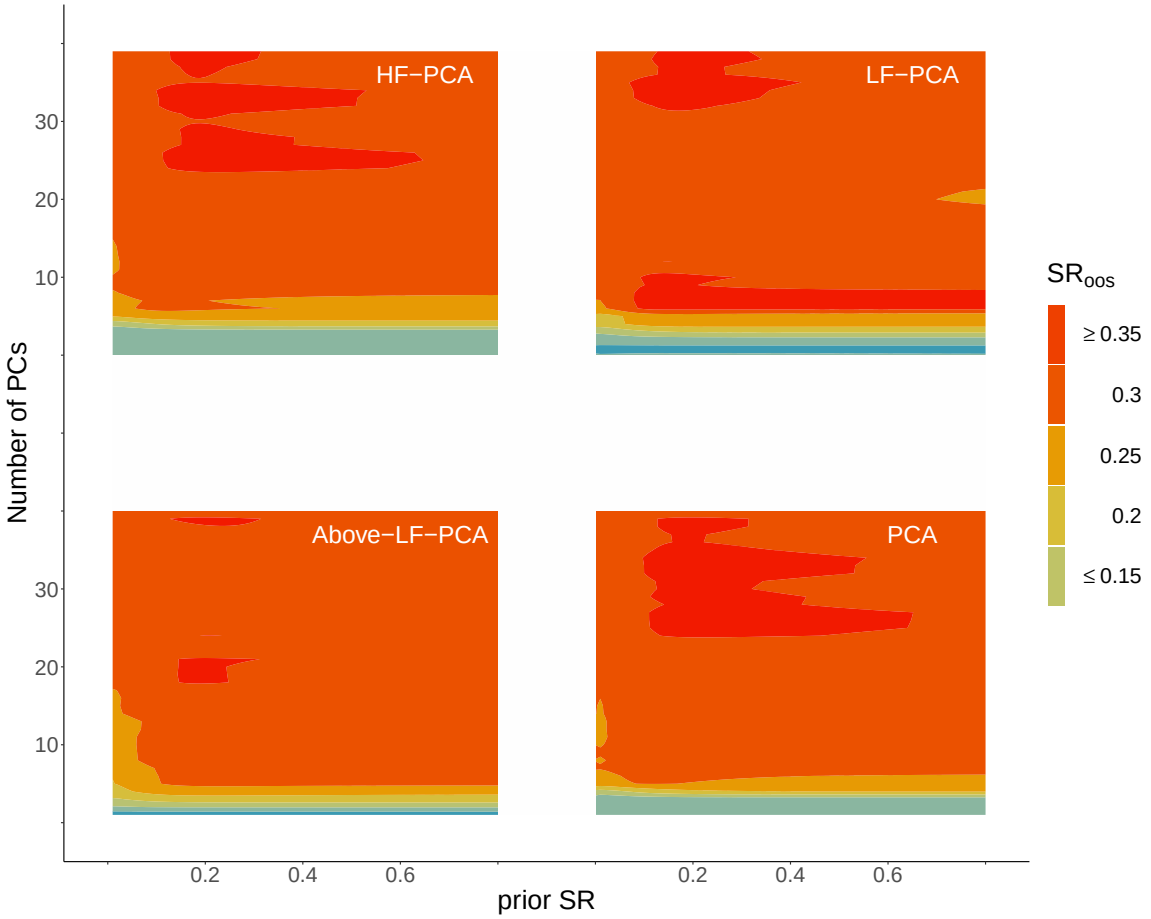


Figure A8: OOS Sharpe ratio of 39 test assets

This graph plots the heat-maps of OOS Sharpe ratios of HF-PCA, LF-PCA, Above-LF-PCA and PCA in the cross-section of 39 long-short portfolios. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. I include the PCs into the SDF based on their ability to explain time-series variations of 39 long-short portfolios.

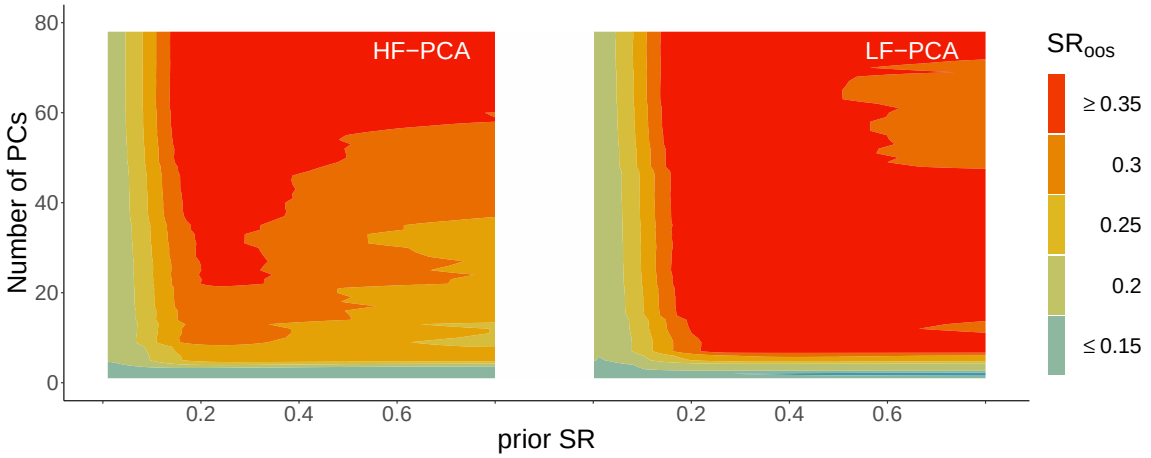


Figure A9: Robustness Check: OOS Sharpe ratio of 78 test assets, $\tau^{LF} \in [24, 120]$

This graph plots the heat-maps of OOS Sharpe ratios of HF-PCA and LF-PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. I include the PCs into the SDF based on their ability to explain time-series variations. The LF frequency interval is defined as $\tau^{LF} \in [24, 120]$.

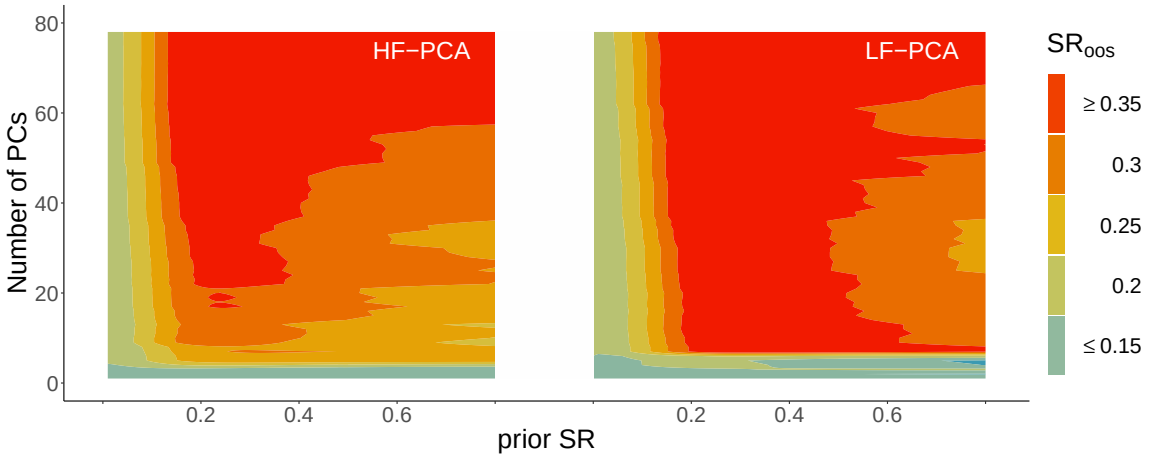


Figure A10: Robustness Check: OOS Sharpe ratio of 78 test assets, $\tau^{LF} \in [32, 64]$

This graph plots the heat-maps of OOS Sharpe ratios of HF-PCA and LF-PCA in the cross-section of 78 test assets. In each panel, the x-axis denotes the prior Sharpe ratio of the factor model, while the y-axis is the number of PCs included in the SDF. In addition, different colors represent different OOS Sharpe ratios. I include the PCs into the SDF based on their ability to explain time-series variations. The LF frequency interval is defined as $\tau^{LF} \in [32, 64]$.