Asset-Pricing Factors with Economic Targets*

Svetlana Bryzgalova† Victor DeMiguel‡ Sicong Li§ Markus Pelger¶

December 2022

Abstract

We propose a method to estimate latent asset-pricing factors that incorporates economically motivated targets for both cross-sectional and time-series properties of the factors. Cross-sectional targets may capture the shape of loadings (monotonicity of expected returns across characteristic-sorted portfolios) or the pricing span of exogenous state variables (macroeconomic innovations or intermediary-based risk factors). Time-series targets may capture overall expected returns or mispricing relative to a benchmark reduced-form model. Using a large-scale set of assets, we show that these targets nudge risk factors to better span the pricing kernel, leading to substantially higher Sharpe ratios and lower pricing errors than conventional approaches.

Keywords: Cross-section of asset returns, portfolio sorts, principal component analysis, shape restrictions, factor identification, latent factors.
JEL classification: C14, C52, C58, G12

*We thank Alberto Martin-Utrera and seminar participants at INFORMS Annual Meeting and Conference on Computational and Financial Econometrics.

†London Business School and CEPR, sbryzgalova@london.edu
‡London Business School, avmiguel@london.edu
§London Business School, ali@london.edu
¶Stanford University, Department of Management Science & Engineering, mpelger@stanford.edu.

1 Introduction

The fundamental insight of asset pricing is that expected returns are driven by asset exposure to systematic sources of risk. As a result, the last few decades have seen the development of two often separate strands of literature on the estimation and testing of risk factors. On one hand, there are powerful statistical approaches to recover risk factors from observed asset returns such as Principal Component Analysis (PCA) and its variations.[1] On the other hand, numerous papers propose and test economically motivated factors, relying on both structural and reduced-form models.[2] This paper bridges the gap between these two strands of literature and shows that incorporating economically motivated targets that capture both structural and reduced-form economic insights significantly improves the performance of statistical risk factors.

We propose a general approach to estimate statistical risk factors that aligns their construction and performance along different *economic targets* (restrictions). Our approach generalizes PCA and accommodates different types of cross-sectional and time-series restrictions on the risk-factor span, pricing ability, and loading pattern (shape). Imposing economic restrictions allows us to nudge latent risk factors to better reflect the insights from theoretical and empirical asset pricing, and thus, significantly improve their performance in terms of Sharpe ratios or pricing errors.

We start by incorporating cross-sectional economic targets on the recovery of statistical risk factors, an approach we term cross-sectional target PCA (XS-target-PCA). While our approach allows for a very general formulation of the cross-sectional targets, we focus on two particular instances. First, inspired by the macroeconomic literature that studies the role of shape restrictions in the identification of structural shocks (pioneered by the long-run restrictions of Blanchard and Quah (1989) and the sign restrictions of Uhlig (2005)), we show how to impose intuitive shape restrictions on the risk-factor loadings. The overwhelming majority of portfolios sorted by a single characteristic (decile sorts) exhibit a monotonic pattern in expected returns, which routinely leads to the creation of long-short factors that reflect a significant spread in asset returns (see Fama and French (1993), Patton and Timmermann (2010), Hou, Xue, and Zhang (2018), Jensen, Kelly, and Pedersen (2022)). Driven by this observation, we impose monotonicity in the expected returns of decile-sorted portfolios on the pricing span achieved by latent factors.

Second, since tradable factors are often thought to reflect systematic sources of risk originated by nontradable (and empirically poorly measured) state variables (macroeconomic innovations or shocks to intermediary risk capacity), we can explicitly impose this restriction on factor recovery. Intuitively, the pricing span of risk factors should include that of exogenously given

[1] See Chamberlain and Rothschild (1983), Connor and Korajczyk (1986), Connor and Korajczyk (1988), Kim, Korajczyk, and Neuhierl (2020), Raponi, Robotti, and Zaffaroni (2019), Fan, Liao, and Wang (2016), Kelly, Pruitt, and Su (2019), Lettau and Pelger (2020b)

[2] See Sharpe (1964) and Lintner (1965), Chen, Roll, and Ross (1986), Jagannathan and Wang (1996), Fama and French (1993), Lettau and Ludvigson (2001), Hou, Xue, and Zhang (2015), Fama and French (2015), He, Kelly, and Manela (2017).

sources of risk, captured by their covariance with asset returns. As a result, this nudges latent factors to directly reflect the asset-pricing implications of state variables that have been identified by economic theory.

We also consider incorporating time-series economic targets on the recovery of statistical risk factors to exploit information in asset risk premia, an approach we term time-series target PCA (TS-Target-PCA). We again allow for a general formulation of the targets, and focus on two particular cases. Risk factors should not only capture the source of systematic time-series variation in asset returns, but, consistent with the strict version of the Arbitrage Pricing Theory (APT), also carry risk premia that explain cross-sectional differences in expected returns. The risk-premia PCA of Lettau and Pelger (2020b) is designed to achieve this property and has been shown to significantly improve the performance of the resulting latent factors in pricing the cross-section of asset returns. Our general framework nests the setup of Lettau and Pelger (2020b) as a particular type of the time-series target that can be imposed on latent factors.

Because in general APT allows for bounded asset-pricing errors (Huberman (1982), Chamberlain and Rothschild (1983), Uppal, Zaffaroni, and Zviadadze (2021)), it might be useful to nudge risk factors towards spanning the pricing errors relative to a benchmark (imperfect) asset-pricing model. To do this, we formulate a time-series target that incorporates alphas from the regression of asset returns on a popular risk factor model such as the five-factor model of Fama and French (2015) (FF5), and encourage statistical risk factors to capture the part of expected returns that the popular model *fails to explain*.

We evaluate the empirical performance of our approach on a large-scale dataset including excess returns of 370 decile portfolios constructed from single sorts on 37 firm characteristics. Note that our approach is not designed to deliver the single most efficient way of identifying factors that explain the cross-section of expected returns. Instead, we present a general framework that allows the researcher to incorporate her ex-ante beliefs about different patterns of returns into the recovery of the latent risk factors and the underlying Stochastic Discount Factor (SDF). As a result, whether or not our approach delivers superior cross-sectional performance compared to standard PCA or other approaches, fully depends on the "quality" of the economic restrictions, and how informative they are about the underlying risk factors.

Our empirical study reveals five key insights. First, we find that imposing cross-sectional shape restrictions (consistent with the monotonic patterns in expected returns of decile-sorted portfolios) significantly improves the empirical performance of latent asset-pricing factors, which achieve higher out-of-sample Sharpe ratio and lower pricing errors compared to standard PCA factors. Second, incorporating either of the time-series targets in factor estimation (mean returns or alphas) also leads to a significant improvement in empirical performance. As expected, each of the time-series targets imposes its own restriction on the risk factors: while the TS-mean-PCA factors yield low pricing errors for characteristic-sorted portfolios, the TS-alpha-PCA factors better capture the average returns of portfolios that are notoriously challenging for the FF5 factor model (e.g., reversal). Third, there are benefits to imposing both cross-sectional and time-series targets on factor recovery. In particular, imposing both shape restrictions on the factor loading pattern as well as the APT-driven time-series pricing restrictions, leads to risk factors that strongly reflect the underlying long and short patterns of asset returns, while preserving high Sharpe ratios and low pricing errors. Our fourth empirical observation is that nudging latent factors to span popular macroeconomic state variables does not help to explain a large cross-section of portfolio returns. This may be due to the fact that most of the pricing ability of macroeconomic risk factors is likely to manifest over the long horizon, and hence, does not provide a sizable improvement to the identification of risk factors within a relatively short investment horizon (see Dew-Becker and Giglio (2016), Neuhierl and Varneskov (2021)). Fifth, we find that the improvement in the cross-sectional explanatory power of latent factors obtained using our cross-sectional and time-series targets does not come at the expense of a reduced time-series explanatory power.

Our paper contributes to an emerging literature that uses new econometric techniques in asset pricing for high-dimensional data, including machine-learning methods. Perhaps the closest work to ours is that of Lettau and Pelger (2020b), who introduce an APT-driven requirement for risk factors to span expected returns of the portfolios (in addition to their time-series variation). Our general framework nests RP-PCA as a particular case of time-series target. In fact, RP-PCA directly corresponds to imposing the time-series mean target on the estimation of latent factors, which we consider as one of the empirical applications. Our approach is more general, because it not only allows one to include different types of the time-series targets, but also impose additional cross-sectional restrictions/nudges on the shape and structure of latent factors. Our XS-Target-PCA is grounded in formal asymptotic theory, which is derived under general assumptions in Duan, Pelger, and Xiong (2022). While Duan et al. (2022) study the general problem of estimating a latent factor model with missing observations by optimally using the information from auxiliary panel data sets, our focus is on including economic cross-sectional and times-series targets for estimating asset pricing factors.

Approximate factor models have been a leading framework to explain a wide cross-section of asset returns (see Ross (1976) and Chamberlain and Rothschild (1983)). While the conditions and methods for strong factors are well established in the existing literature (Connor and Korajczyk (1986), Stock and Watson (2002), Bai (2003)), inference on semi-strong and weak factors is a challenge. Onatski (2012) studies PCA-based estimation within the context of weak factors, and shows that unless a weak factor explains a sufficient amount of the time-series variation in the data, it cannot be statistically detected. Our approach, based on cross-sectional and time-series targets, provides a new source of information for identifying the impact of latent factors. Similar to valid economic constraints, the targets can significantly reduce estimation uncertainty and improve factor identification.

There is a large and actively growing literature that studies the high-dimensional nature of asset returns and its underlying structure. Freyberger, Neuhierl, and Weber (2020) and Feng, Giglio, and Xiu (2020a) employ factor selection using L^1-norm penalties and assume sparsity in the underlying SDF exposure. Similar to Kozak, Nagel, and Santosh (2020), our approach does not assume that the SDF is sparse in any explicit economic signal. Instead, we rely on the cross-sectional and time-series targets to encode information of economically important restrictions/priors, and use the PCA framework to identify these factors.

Using latent factors in asset pricing, especially due to prevalent model misspecification, recently became one of the most popular approaches to recovering the span of the SDF. Giglio and Xiu (2021) use PCA to identify factors that price the cross-section of expected returns, and estimate the price of risk in the presence of model misspecification. Dello Preite, Uppal, Zaffaroni, and Zviadadze (2022) complement this approach by also considering asset-specific risk. Kelly et al. (2019) propose the Instrumented-PCA (IPCA) approach, which is closely related to the projected-PCA of Fan et al. (2016). Assuming that individual stocks have factor loadings that are linear in firm characteristics, they extract latent factors via a version of the PCA, applied to characteristic-managed portfolios. Our framework can be used at the level of managed portfolios, and since it allows to efficiently encode the researcher's beliefs regarding cross-sectional and time-series properties of the latent factors, nothing precludes using it as part of the IPCA procedure as well. If the underlying economic restrictions are informative about the systematic sources of risk, using them might lead to significant efficiency gains and better identification of risk factors. This in turn, may lead to superior empirical performance at the level of both managed portfolios and individual stocks. Since the main objective of our paper is to demonstrate how different economic priors can be efficiently encoded in risk-factor estimation, we leave this empirical exercise for future work.

More generally, our work is also related to recent applications of machine-learning techniques in asset pricing. Freyberger, Neuhierl, and Weber (2016), Gu, Kelly, and Xiu (2018), Chen, Pelger, and Zhu (2020), Bianchi, Buchner, and Tamoni (2022) demonstrate that using machine-learning techniques leads to a substantial improvement in the prediction of stock and bond returns. Feng, Polson, and Xu (2020b) use deep learning based on characteristics and macroeconomic predictors to estimate latent factors and incorporate them into a model that already contains observed ones. Gu, Kelly, and Xiu (2021) extend the IPCA by allowing the factor exposures to be a nonlinear function of the covariates, and estimate these exposures using autoencoder neural networks. Bali, Goyal, Huang, Jiang, and Wen (2022) demonstrate that bond return predictability with firm characteristics drastically improves when one imposes a theoretically motivated model for asset returns (e.g., Merton model), compared to a simple yet flexible reduced-form approach. Similar to Kozak et al. (2020) and Bryzgalova, Huang, and Julliard (2022), we impose economic restrictions to recover the underlying SDF. Our restrictions (targets) are not limited to the use of macroeconomic variables or imposing certain properties on the overall time-series features of the underlying latent factors.

Importantly, we provide a general framework to encode various economic restrictions, stemming from stylized theoretical models or empirical intuition of researchers, into a latent factor model.

The rest of the paper is organized as follows. Section 2 describes the proposed cross-sectional time-series PCA approach. Section 3 evaluates the performance of the proposed approach using a high-dimensional dataset of 370 decile-sorted portfolios, and Section 4 concludes.

2 Factor models with economic targets

2.1 Factor model

Assume that excess returns follow a standard approximate factor model: Their dynamics are driven by a systematic component, captured by K risk factors, and a nonsystematic (idiosyncratic) one accounting for the asset-specific risk (which can be weakly dependent). We observe excess returns of N assets over T time periods:

$$R_{n,t} = F_t \beta_n^\top + e_{n,t},$$

or, in matrix notation,

$$\underbrace{R}_{T\times N} = \underbrace{F}_{T\times K} \underbrace{\beta^\top}_{K\times N} + \underbrace{e}_{T\times N}, \tag{1}$$

where the unknown factors F and loadings (or betas) β have to be estimated. We allow for large dimensional panels, that is both the cross-sectional dimension N and time-series dimension T, can be large. We denote by Σ_R and Σ_F the variance-covariance matrices of returns and factors, and by Σ_e the covariance matrix of residuals. Assuming that factor returns and errors are uncorrelated, the covariance matrix of asset returns can be decomposed into systematic and idiosyncratic parts:

$$\Sigma_R = \beta \Sigma_F \beta^\top + \Sigma_e. \tag{2}$$

If factors are sufficiently strong, that is, they explain enough time-series variation in the panel, we can estimate F via the Principal Component Analysis (PCA). The core assumption is that the eigenvalues of the systematic part are asymptotically much larger than the eigenvalues from the residuals. This is the case if many assets in the panel have non-vanishing loadings on the factors and the variance of the factors is sufficiently large. Formally, a conventional PCA estimator minimizes the following objective function:

$$\text{PCA:} \quad \hat{F}_{\text{PCA}}, \hat{\beta}_{\text{PCA}} = \underset{F,\beta}{\operatorname{argmin}} \frac{1}{NT} \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} \left(\left(R_{t,n} - \bar{R}_n \right) - \left(F_t - \bar{F} \right) \beta_n^\top \right)^2$$

$$= \underset{F,\beta}{\operatorname{argmin}} \frac{1}{NT} \left\| \left(R - \bar{R}\right) - \left(F - \bar{F}\right) \beta^\top \right\|_F^2,$$

where $\bar{R}_n$ and $\bar{F}$ denote the sample mean of the returns and factors, respectively, and $\|\cdot\|_F$ is the Frobenius norm. Under the standard normalization that the loadings are orthonormal $\frac{1}{N}\hat{\beta}_{\text{PCA}}^\top \hat{\beta}_{\text{PCA}} = I_K$, the PCA estimator follows from a spectral decomposition of the sample covariance matrix of returns. The estimated loadings equal the $N \times K$ matrix of eigenvectors scaled by $\sqrt{N}$ of the K largest eigenvalues of $\hat{\Sigma}_R$. The factors then follow from a cross-sectional regression on the loadings as $\hat{F}_{\text{PCA}} = R\hat{\beta}_{\text{PCA}} \left(\hat{\beta}_{\text{PCA}}^\top \hat{\beta}_{\text{PCA}}\right)^{-1}$. In other words, latent factors are linear combinations (portfolios) of the original returns R, with the following portfolio weights $\hat{\omega}_{\text{PCA}} \in \mathbb{R}^{N\times K}$:

$$\hat{F}_{\text{PCA}} = \frac{1}{N} R\, \hat{\omega}_{\text{PCA}} \qquad \hat{\omega}_{\text{PCA}} = \hat{\beta}_{\text{PCA}}.$$

As long as factors are uncorrelated, the factor weights and conventional OLS loading estimators are the same, but we will distinguish between them for our more general estimation approaches. Following the seminal contribution of Bai (2003), formal properties of the PCA-based estimation of latent factors and loadings have been established in various empirical settings.

Conventional PCA-based estimation of latent factors is known to suffer from significant drawbacks in many asset pricing frameworks. Naturally, its objective is to find latent factors that explain as much variation as possible for a given set of basis assets, R. However, to do so it only leverages information in the second moment (variance-covariance matrix), and ignores all the other information, e.g., other moments (including the mean), additional sources of information, and potential economic restrictions. Furthermore, it puts the same weight on all basis assets. As a result, economically important factors that explain the variation in risk premium, that is, the mean of the asset returns, yet affect only a smaller subset of the cross-section and/or have a low variance that might not be detectable with the PCA.

A popular solution, suggested in the literature, has been to include additional information in the estimation, whenever the covariance signal is not sufficient to detect weak factors. This additional information can be based on economic assumptions or priors. In particular, the Risk-Premium-PCA (RP-PCA) of Lettau and Pelger (2020b) leverages the Arbitrage Pricing Theory (APT) and the information included in the risk premium in order to boost the signal of weak factors. Under APT, the mean of asset returns is approximately explained by the factor means and their risk exposure, that is, $E[R] = \beta[F]$. As a result, RP-PCA adds a penalty to the conventional PCA estimator to find factors that explain the most variation subject to the APT condition. This additional information allows to detect factors that are weak in terms of the explained time-series variation, yet important for the estimation of risk premia.

In this paper, we propose a general class of estimators that allow to leverage information from both cross-section and time-series of returns and additional data (e.g., macroeconomic innovations) in recovery of the asset pricing factors, that includes Lettau and Pelger (2020b) as a particular case. We show that imposing economically motivated restrictions can boost the signals of weak factors that are quite challenging to identify with conventional reduced-form methods. In particular, we focus on the information content of economic fundamentals (which impact should be reflected by latent asset pricing factors), shape restrictions in the patterns of asset loadings on risk factors, and cross-sectional pricing ability of the factors, driven by average returns or alphas relative to the benchmark reduced-form models. While imposing these economically motivated restrictions, we nudge the conventional factor estimation towards certain time-series (TS) or cross-sectional (XS) targets, yielding the new estimation framework, XS-TS-Target PCA.

2.2 Cross-sectional targets: XS-Target PCA

First, we show how to include cross-sectional targets in the latent factor estimation. XS-Target PCA adds a penalty to estimate factors that align with an economically motivated cross-sectional structure. In the general case, we model cross-sectional target as a matrix $\Lambda^{\text{XS}} \in \mathbb{R}^{N\times L}$, such that it highly correlates with the matrix of factor loadings β. Formally, this requires a subspace spanned by the loadings to be included in the space spanned by the target. If factor loadings should reflect a certain pattern (e.g., shape restrictions like monotonicity, or capturing the price impact of exogenously given risk factors, such as consumption or intermediary capital ratio), nudging β to explain the structure in Λ could significantly improve factor recovery.

In order to provide some intuition behind the estimator, we assume that the scaled matrix Λ^{XS} is normalized to be orthonormal, that is $\frac{1}{N}(\Lambda^{\text{XS}})^\top \Lambda^{\text{XS}} = I_L$. This condition can easily be relaxed, and does not impose any significant constraint on the general approach to estimation. However, it simplifies the exposition, and, therefore, is used for illustrative purposes. The cross-sectional target implies the following set of managed target portfolios:

$$\underbrace{R^{\text{XS}}}_{T\times L} = \frac{1}{N} R \Lambda^{\text{XS}}.$$

Conceptually, we would like the factor model to explain both the basis assets R and the target portfolios R^{XS}. Hence, the estimator solves the following optimization problem

$$\underset{F,\beta}{\text{argmin}}\, \frac{1}{NT} \sum_{n=1}^{N} \sum_{t=1}^{T} \left(R_{t,n} - F_t \beta_n^\top \right)^2 + \gamma_{\text{XS}} \left(\frac{1}{LT} \sum_{l=1}^{L} \sum_{t=1}^{T} \left(R_{t,l}^{\text{XS}} - F_t (\beta_l^{\text{XS}})^\top \right)^2 \right)$$

with $(\beta^{\text{XS}})^\top = \beta^\top \Lambda^{\text{XS}}$ and the penalty γ_{XS}. In the special case of $\gamma_{\text{XS}} = 0$, we recover a conventional PCA estimator. Naturally, increasing γ_{XS} pushes latent factors to better explain target portfolios.

This estimator, of course, involves a projection problem. To address it, we relax the normalization that Λ^{XS} is orthonormal and define the *cross-sectional projection matrix* $P_\Lambda^{\text{XS}} \in \mathbb{R}^{N\times N}$

$$P_\Lambda^{\text{XS}} = \Lambda^{\text{XS}} \left((\Lambda^{\text{XS}})^\top \Lambda^{\text{XS}} \right)^{-1} (\Lambda^{\text{XS}})^\top, \tag{3}$$

which projects N dimensional vectors onto the column space of Λ^{XS}. XS-Target-PCA solves the following objective

$$\text{XS-Target-PCA:} \quad \hat{F}_{\text{XS}}, \hat{\beta}_{\text{XS}} = \underset{F,\beta}{\operatorname{argmin}} \; \underbrace{\frac{1}{NT} \left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{variation of returns}}} + \gamma_{\text{XS}} \underbrace{\frac{1}{LT} \left\| RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}} \right\|_F^2}_{\substack{\textbf{XS target}\text{: unexplained variation} \\ \text{of } P_\Lambda\text{-projected returns}}} . \tag{4}$$

Note that $\left\| RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}} \right\|_F^2 = \sum_{l=1}^L \sum_{t=1}^T \left(R_{t,l}^{\text{XS}} - F_t (\beta_l^{\text{XS}})^\top \right)^2$ if we normalize Λ^{XS}.

Hence, XS-Target-PCA nudges the estimator to fit certain target assets. If the target assets have an informative economic structure, which is weak in the original panel R, then setting γ_{XS} sufficiently large allows to detect these weak factors. Note that the dimensions of the panel of basis assets and target assets can be very different, but can be accounted for by an appropriate choice of γ_{XS}.

The following proposition shows that the latent-factor model that optimizes the XS-Target-PCA objective in (4) can be obtained by applying PCA to a transformed second moment matrix.

Proposition 1 *The factors that optimize the XS-Target-PCA objective are obtained by applying PCA to the $T \times T$ matrix*

$$\frac{1}{NT} R \left(I_N + \gamma_{XS} \frac{N}{L} P_\Lambda^{XS} \right) R^\top. \tag{5}$$

Specifically, the estimated factor returns, $\hat{F}_{XS}$, are the first K eigenvectors of matrix (5)*, and the estimated factor weights and loadings are $\hat{\omega}_{XS} = \hat{\beta}_{XS} = R^\top \hat{F}_{XS}$.*

In our empirical analysis we consider two types of cross-sectional targets. The first target is economically motivated shape restrictions in the structure of the loadings. Most asset pricing characteristics are motivated by a univariate analysis that shows monotonic patterns in the risk premia, aligned with certain firm characteristics (that is, single-sorted portfolios tend to produce a spread in expected returns). Therefore, we expect the loadings of asset-pricing factors to also reflect an underlying monotonic structure. We incorporate this economic prior by a linearly increasing pattern within the deciles sorts of a large panel of portfolios with many characteristics. More specifically, in our empirical analysis we consider $N = 370$ decile-sorted portfolios for 37 different characteristics. Our target matrix, therefore, has the target dimension $L = 37$. Each column targets a specific characteristic with a linear structure centered at zero for the ten deciles of a specific characteristic and zero otherwise. This structure is illustrated in Figure 1.

The monotonic shape restriction has two effects: First, it includes the additional information that characteristic sorts are expected to have a monotonic risk premium structure. Hence, it imposes a structure that is related to mean returns. Intuitively, the target-managed portfolios related to the shape target are univariate long-short factors. Thus, the XS-Target-PCA extracts factors that do not only explain the panel of $N = 370$, but also specifically long-short univariate risk premia. Second, penalizing the shape of the loadings also nudges the loadings and factor weights to tilt towards this structure. Hence, as we show empirically, XS-Target-PCA factors with a shape target are long-short factors within the characteristic deciles, while conventional PCA factors do not possess this structure. This benefits the interpretability of the statistical factors. While our analysis focuses on monotonic shapes, it is straightforward to include alternative shapes to capture for example convexity effects.

The second target is based on economic fundamentals. More specifically, our factors are nudged to be more correlated with fundamental macroeconomic innovations. This target is based on the theoretical insight of the intertemporal CAPM of Merton (1973) that innovations to macroeconomic variables that affect the marginal utility of consumption of a representative investor represent systematic risks that should be priced. Motivated by this theoretical insight, we consider the target matrix Λ^{XS} of correlations of the basis assets with $L = 7$ important macroeconomic time series. Given a panel of L macroeconomic time series Y^{macro}, we obtain the macroeconomic target as

$$\Lambda^{\text{XS}} = RY^{\text{macro}} \left((Y^{\text{macro}})^\top Y^{\text{macro}} \right)^{-1} \tag{6}$$

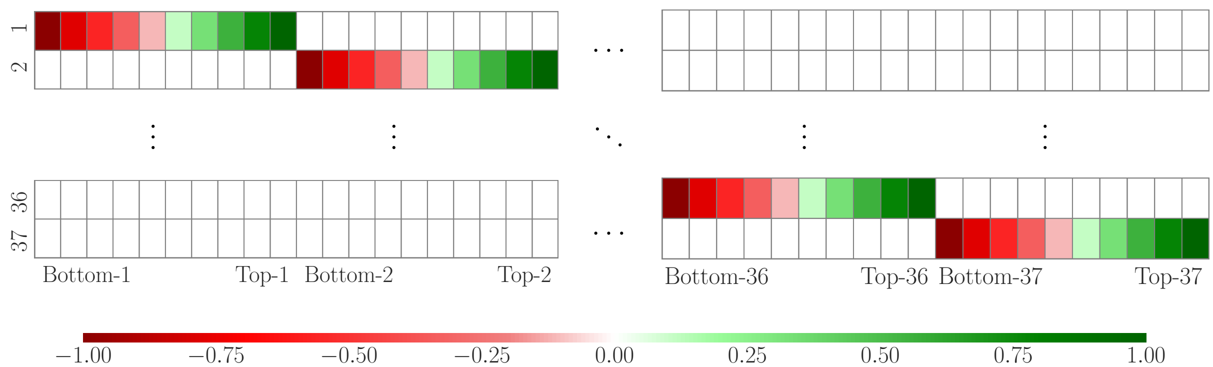
Hence, the target mimicking portfolios R^{XS} correspond to Fama-MacBeth type macroeconomic mimicking factor portfolios. The XS penalty forces the extracted factors to explain both the basis assets and these macroeconomic mimicking portfolios. An appropriate choice of the penalty allows to take advantage of the weak correlation between returns and macroeconomic innovations. Hence, if these innovations in economic fundamentals contain cross-sectional asset pricing information, XS-Target-PCA will improve upon a conventional PCA estimator.

2.3 Time-series targets: TS-Target-PCA

The time-series targets take advantage of return moments in addition of the second moment. The conventional PCA estimator only uses the second moment to extract latent factors. However, if the cross-sectional structure in other time-series moments is informative for the cross-section, then they should also be included. The most prominent case is the first moment. The arbitrage pricing theory (APT) implies that factors that drive the systematic time-series variation of asset returns should also explain the cross-section of expected returns. This additional moment allows to estimate factors that have a too-weak covariance signal to be detected by conventional PCA.

Figure 1: Cross-sectional shape target

This figure illustrates the cross-sectional shape target $\Lambda^{\mathrm{XS}} \in \mathbb{R}^{370\times 37}$ for 37 different characteristics. Each row of the figure corresponds to one column of Λ^{XS} and thus targets one characteristic with a linear structure centered at zero for its ten decile portfolios and zero otherwise. The vertical axis depicts the 37 characteristics, and the horizontal axis depicts the 370 decile portfolios of the characteristics. Positive (negative) values in Λ^{XS} are illustrated in green (red). Zeros are illustrated in white.



Lettau and Pelger (2020b) use a time-series target that nudges factors to explain average asset returns. In particular, their Risk-Premium PCA (RP-PCA) estimator augments the PCA objective with a time-series target that penalizes cross-sectional pricing errors:

$$\text{RP-PCA:} \quad \hat{F}_{\mathrm{RP}}, \hat{\beta}_{\mathrm{RP}} = \underset{F,\beta}{\operatorname{argmin}} \quad \underbrace{\frac{1}{NT}\left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{time-series variation}}} + \gamma_{\mathrm{RP}} \underbrace{\frac{1}{N}\left\| \bar{R} - \frac{1}{T}\bar{F}\beta^\top \right\|_F^2}_{\substack{\textbf{TS target}\text{: cross-sectional} \\ \text{pricing error}}}, \tag{7}$$

where γ_{RP} is risk-premium penalty parameter.

Our TS-Target-PCA generalizes RP-PCA to leverage the cross-sectional information in general time-series moments. We allow for J general target moments that are of the form $\frac{1}{T}RG^{\mathrm{TS}}$ for the target TS matrix $G^{\mathrm{TS}} \in \mathbb{R}^{T\times J}$. RP-PCA is the special case of setting $G^{\mathrm{TS}} = \mathbf{1}_T$, where $\mathbf{1}_T$ is the T-dimensional vector of ones. The cross-sectional span of the loadings β is nudged to be aligned with the cross-sectional span of the target moments $\frac{1}{T}RG^{\mathrm{TS}}$. For this purpose, we introduce the *time-series projection matrix* $P_G^{\mathrm{TS}} = G^{\mathrm{TS}}((G^{\mathrm{TS}})^\top G^{\mathrm{TS}})^{-1}(G^{\mathrm{TS}})^\top$, which projects T-dimensional vectors onto the column space of G^{TS}. TS-Target-PCA estimates latent factors using the following objective

$$\text{TS-Target-PCA:} \quad \underset{F,\beta}{\operatorname{argmin}} \quad \underbrace{\frac{1}{NT}\left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{time-series variation}}} + \gamma_{\mathrm{TS}} \underbrace{\frac{1}{NT}\left\| P_G^{\mathrm{TS}}R - P_G^{\mathrm{TS}}F\beta^\top \right\|_F^2}_{\substack{\textbf{TS target}\text{: moment error for} \\ P_G^{\mathrm{TS}}\text{-projected returns}}}, \tag{8}$$

where γ_{TS} is the penalty parameter for the time-series target.

The factors that optimize the TS-Target-PCA objective in (8) can be estimated by applying a simple PCA to a transformed $N \times N$ matrix as shown in the following proposition.

Proposition 2 *The latent factors that optimizes the TS-PCA objective can be estimated by applying PCA to the $N \times N$ matrix*

$$\frac{1}{NT} R^\top \left(I_T + \gamma_{TS} P_G^{TS} \right) R \tag{9}$$

Specifically, the estimated factor weights $\hat{\omega}_{TS}$ are the first K eigenvectors of matrix (A11)*, and the factor returns are $\hat{F}_{TS} = R\hat{\omega}_{TS}$.*

As the factors are not necessarily orthogonal, the factor weights $\hat{\omega}_{\text{TS}}$ are not identical to the OLS loadings $\hat{\beta}_{\text{TS}}$. However, the loadings $\hat{\beta}_{\text{TS}}$ estimated by an OLS regression of $\hat{F}_{\text{TS}}$ on R are consistent estimators of the loadings β_{TS}.

We consider two types of time-series targets. First, we include the important case of the mean, that is, we use RP-PCA. Lettau and Pelger (2020b) have shown that RP-PCA factors substantially outperform PCA factors for asset pricing. Second, we introduce Alpha-PCA, that is, we find factors that explain a cross-section of target alphas. The goal is to extract factors that explain assets that are mispriced by a candidate asset pricing model. This target makes sense, as the mean of the cross-section can be dominated by "easy" factors. Our penalty nudges the estimator to explore hard-to-price assets and put more weight on the weak factors in the alpha portfolios.

In more detail, we calculate cross-sectional pricing errors relative to a set of candidate reference factors F^{ref}. In our empirical analysis, we consider the five factors of Fama and French (2015). In this case, the projection matrix should be based on the following matrix:

$$G^{\text{TS}} = \left(I_T - \tilde{F}^{\text{ref}} \left((\tilde{F}^{\text{ref}})^\top \tilde{F}^{\text{ref}} \right)^{-1} (F^{\text{ref}})^\top \right) \mathbf{1}_T, \tag{10}$$

where F^{ref} and $\tilde{F}^{\text{ref}}$ are the predetermined reference factors and their demeaned counterparts, respectively. Hence, applying G^{TS} to the asset returns R yields $\frac{1}{T} R^\top G^{\text{TS}} = \alpha^{\text{ref}}$, where $\alpha^{\text{ref}} \in \mathbb{R}^N$ is the vector of pricing errors (alphas) of the assets with respect to the reference factors. As a result, TS-Target-PCA penalty nudges the latent factors to the carry risk premia not explained by the reference factors F^{ref}.

2.4 Combining cross-sectional and time-series targets: XS-TS-Target PCA

The cross-sectional and time-series targets can capture different economic information. Therefore, we combine both objectives in our XS-TS-Target PCA. This joint estimator has two penalties to target both the cross-sectional and the time-series structure simultaneously. For example, we can combine the monotonic shape constraint of the loadings with the risk-premium penalty for the factors. As we will show, these two additional information sets are non-redundant, and the combined estimator outperforms using only one of the targets.

The XS-TS-Target PCA obtains factors by solving the following problem:

$$\begin{aligned} \underset{F,\beta}{\operatorname{argmin}} \; & \underbrace{\frac{1}{NT}\left\|R - F\beta^\top\right\|_F^2}_{\substack{\text{unexplained}\\ \text{time-series variation}}} + \gamma_{\text{XS}} \underbrace{\frac{1}{LT}\left\|RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}}\right\|_F^2}_{\substack{\textbf{XS target}\text{: unexplained variation}\\ \text{of } P_\Lambda^{\text{XS}}\text{-projected returns}}} \\ & + \gamma_{\text{TS}} \underbrace{\frac{1}{NT}\left\|P_G^{\text{TS}}R - P_G^{\text{TS}}F\beta^\top\right\|_F^2}_{\substack{\textbf{TS target}\text{: moment error for}\\ P_G^{\text{TS}}\text{-projected returns}}} + \gamma_{\text{XS}} \cdot \gamma_{\text{TS}} \underbrace{\frac{1}{LT}\left\|P_G^{\text{TS}}RP_\Lambda^{\text{XS}} - P_G^{\text{TS}}F\beta^\top P_\Lambda^{\text{XS}}\right\|_F^2}_{\substack{\textbf{XS-TS target}\text{: } P_G^{\text{TS}}\text{-moment error}\\ \text{of } P_\Lambda^{\text{XS}}\text{-projected returns}}}, \end{aligned} \tag{11}$$

The first term is the unexplained time variation, which is the usual PCA-type objective. The second term is the cross-sectional penalty of XS-Target-PCA and captures the explained variation in managed target portfolios R^{XS}. The third term captures the errors in target moments of the full panel and corresponds to the penalty of TS-Target-PCA. The last term combines the cross-sectional and time-series targets. It penalizes the error in time-series moments of the managed target portfolios R^{XS}.

The general XS-TS-Target PCA objective can also be estimated by applying PCA to a transformed matrix, as explained by the next proposition.

Proposition 3 *The latent factors that optimize the XS-TS-PCA objective in* (11) *are estimated by applying PCA to the $N \times N$ matrix*

$$\frac{1}{NT}\left(I_N + \tilde{\gamma}_{XS}P_\Lambda^{XS}\right) R^\top \left(I_T + \gamma_{TS}P_G^{TS}\right) R \left(I_N + \tilde{\gamma}_{XS}P_\Lambda^{XS}\right), \tag{12}$$

where $\tilde{\gamma}_{XS} = \sqrt{\gamma_{XS}\frac{N}{L} + 1} - 1$. Specifically, let $Q_K \in \mathbb{R}^{N\times K}$ be matrix whose columns contain the first K eigenvectors of matrix (12)*, then the estimated factor are $\hat{F}_{XS\text{-}TS} = R\hat{\omega}^{TS\text{-}XS}$ with the estimated factor weights*

$$\hat{\omega}^{TS\text{-}XS} = \left(\frac{1}{\tilde{\gamma}_{XS}+1} I_N + \frac{\tilde{\gamma}_{XS}}{\tilde{\gamma}_{XS}+1} P_\Lambda\right) Q_K. \tag{13}$$

Given the estimated $\hat{F}_{\text{XS-TS}}$, we obtain consistent estimates of the loadings with an OLS regression of the factors on the returns. Proposition A.1 in the Appendix collects the different estimators and provides equivalent ways to estimate them. Similar to a conventional PCA, the singular values can be either estimated from a transformed $N \times N$ or transformed $T \times T$ matrix. The resulting estimator are numerically identical. The main text presents the results for the $N \times N$ spectral decomposition as this is closer to conventional estimation of PCA. After appropriate transformations, the spectral decomposition yields estimator of the latent factors and the loadings. If we include a TS moment target, the loadings based on the spectral decomposition correspond to weighted regression of the factors on the returns, instead of an OLS regression. In our empirical analysis we only extract the latent factors with the modified PCA estimation, and obtain the loadings from a regular OLS regression. These loadings are consistent estimators and empirically very close to the weighted regression loadings.[3] We present the OLS loadings as they allow us to focus on the effect on the factor estimation and correspond to the common practice in asset pricing.

In summary, the XS-TS-Target PCA, and hence also its special cases XS-Target PCA and TS-Target PCA, consists of the following steps:

1. We specify the cross-sectional mapping Λ^{XS} and the time-series mapping G^{TS}. Then, we obtain the corresponding projection matrices P^{XS}_{Λ} and P^{TS}_{G}, respectively.

2. For a given choice of the penalties γ_{XS} and γ_{TS}, we obtain the K latent factors $\hat{F}_{\text{XS-TS}}$ by applying PCA to matrix (12) as outlined in Proposition 3. The special cases XS-Target-PCA and TS-Target-PCA use only one non-zero penalty.

3. We estimate the loadings $\hat{\beta}_{\text{XS-TS}}$ with an OLS time-series regression

$$R_i = \alpha_i + \hat{F}_{\text{XS-TS}}\beta_i^{\top} + e_i, \tag{14}$$

which yields estimates for $\hat{\alpha}$, $\hat{\beta}_{\text{XS-TS}}$ and $\hat{e}$.

2.5 Properties of XS-TS-Target PCA

Including the additional information in the targets can improve the estimation of asset pricing factors. In this section we provide an intuition why it is helpful to include useful targets in the estimation. The formal econometric theory and statistical argument for the XS-Target-PCA is in Duan et al. (2022) and for TS-Target-PCA in Lettau and Pelger (2020a).

A key element in the estimation of latent factors is the factor strength. Intuitively, the strength of a factor in a given panel depends on how much variation is explained by this factor. A factor that affects only a small portion of the assets and/or has little variation is a weak factor. Formally, the strength of factors is the ratio of the systematic eigenvalues due to the factors relative to the non-systematic eigenvalues of the noise. PCA estimation applied to a sample covariance matrix can only consistently estimate latent factors that are sufficiently strong. Weak factors cannot be detected by PCA, even if those factors might carry large risk premia and constitute an important component of the pricing kernel.

[3] The results are available upon request.

The important effect of both the XS and the TS target is that they can increase the strength of weaker factors. While factors might be weak in terms of the variation that they explain in a given panel, they can become stronger in the modified matrix that includes the target information. Simply speaking, the eigenvalues of factors are boosted in the transformed matrix with the target information. This can have different beneficial effects. Weak factors, that cannot be detected from the covariance matrix alone, can now be estimated. Semi-weak factors, that affect small but sufficiently larger portion of the panel, can be estimated with a higher convergence rate. In the case of strong factors, which affect many assets and could be estimated consistently with PCA, the target information can increase the efficiency of the estimation. This is because the XS-TS-Target-PCA can be interpreted as method of moment estimator, and the penalty can be selected to assign efficient weights on the moments.

We will first illustrate the effect of XS-Target-PCA with an example. The formal econometric theory under general assumptions is discussed in Duan et al. (2022). Assume that our panel is described by two factors $R = F\beta + \epsilon$, where (after a possible rotation) the first N_1 assets are only affected by factor one and the last N_2 assets are only affected by factor two:

$$\beta = \begin{pmatrix} \underbrace{\beta_1}_{N_1 \times 1} & \underbrace{0}_{N_1 \times 1} \\ \underbrace{0}_{N_2 \times 1} & \underbrace{\beta_2}_{N_2 \times 1} \end{pmatrix}.$$

Assume that the cross-sectional target selects the first N_1 assets, that is, it is given by the $N \times N_1$ matrix

$$\Lambda_{\text{XS}} = \begin{pmatrix} \underbrace{I_{N_1}}_{N_1 \times N_1} \\ \underbrace{0}_{N_2 \times N_1} \end{pmatrix}.$$

The managed portfolios $R^{\text{XS}} = R\Lambda_{\text{XS}}$ implied by Λ_{XS} simply equal the first N_1 assets, which are only affected by factor one. The key quantity in latent factor estimation are the systematic eigenvalues from the factors given by

$$\begin{aligned} &\frac{1}{NT} F\beta^\top \left(I_N + \gamma_{\text{XS}} \Lambda_{\text{XS}} \Lambda_{\text{XS}}^\top \right) \beta F^\top \\ =&\frac{1}{T} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \left(\begin{pmatrix} \frac{\beta_1^\top \beta_1}{N_1} \frac{N_1}{N} & 0 \\ 0 & \frac{\beta_2^\top \beta_2}{N_2} \frac{N_2}{N} \end{pmatrix} + \gamma_{\text{XS}} \begin{pmatrix} \frac{\beta_1^\top \beta_1}{N_1} \frac{N_1}{N} & 0 \\ 0 & 0 \end{pmatrix} \right) \begin{pmatrix} F_1^\top & F_2^\top \end{pmatrix} \end{aligned}$$

For $N_1/N \to 0$, the panel of returns is not sufficient to identify the first factor F_1 as $\frac{\beta_1^\top \beta_1}{N_1} \frac{N_1}{N} \to 0$ under the usual assumptions, and hence for $\gamma_{\text{XS}} = 0$ the second moment matrix of the loadings is not full rank. In other words, the first factor would not be identified. However, under suitable assumptions selecting $\gamma_{\text{XS}} = r \cdot N/N_1$ results in the combined second moment

$$\begin{pmatrix} \frac{\beta_1^\top \beta_1}{N_1}\frac{N_1}{N} + r \cdot \frac{\beta_1^\top \beta_1}{N_1} & 0 \\ 0 & \frac{\beta_2^\top \beta_2}{N_2}\frac{N_2}{N} \end{pmatrix},$$

which has full rank, and allows to identify both factors. The formal arguments depend on the interplay of the rates of N_1, N_2 and T and the dependency in the noise. The important message is that under realistic assumptions we require a non-trivial weight for γ_{XS} to recover the full factor model. The optimal weight for γ_{XS} depends on the strength of the factors, the dimensionality of the data and the variance of the noise. Duan et al. (2022) provide the optimal values under general assumptions.

Note that this illustrative example captures the empirically relevant case that the target matrix is not sufficient to identify all factors. Hence, simply applying PCA on the managed target portfolios R^{XS} would not recover the full factor model. Hence, the optimal combination of the information in the given panel and the target panel is key.

Even when both factors are strong, that is in our example N_1 and N_2 would both be proportional to N, there can be a benefit of selecting a non-zero γ_{XS}. The noise in the projected data R^{XS} can be lower than in the given panel R. Hence, increasing the weight γ_{XS} can lead to more efficient estimation. As R^{XS} might not identify all factors, we cannot neglect R in the estimation. Note that the use of cross-sectional targets also serves the purpose of alleviating the estimation error by imposing a certain structure on the latent factors. This should be particularly useful in the out-of-sample empirical applications that rely on rolling window estimation. Cross-sectional targets may provide a certain degree of stability by imposing an economically motivated structure on the composition of the factors, hence, improving their empirical stability.

While we used a simple structure for the illustration, the same arguments apply for shape restrictions or targets based on correlations with macroeconomic fundamentals. If the target is not informative, that is, it does not contain information about the factors, then it is optimal to select $\gamma_{\text{XS}} = 0$. Hence, XS-Target-PCA also served as a disciplined approach to diagnose if target information contains useful information in addition to a panel R.

The TS-Target-PCA exploits similar arguments. The formal theory for a mean target is developed in Lettau and Pelger (2020a) and similar arguments would apply to an alpha target. While the intuition is similar to the XS-Target-PCA, the exact formal theory is derived for very weak factors based on random matrix theory.

A crucial object in understanding the properties of TS-Target-PCA is the signal matrix. The signal corresponds essentially to the eigenvalues of the systematic component. Assume that, without loss of generality, we normalize the loadings to be orthonormal, that is, the strength of the factors is completely captured by the factor variance. Intuitively, in the case of a mean target the sample eigenvalues of the systematic component, given by

$$\frac{1}{N}\beta F^{\top}\left(I_T + \gamma_{\text{TS}}\frac{1}{T}\mathbb{1}\mathbb{1}^{\top}\right)F\beta^{\top},$$

should converge to the eigenvalues of the signal matrix

$$\Sigma_F + (1+\gamma_{\text{TS}})\,\mu_F\mu_F^{\top}. \tag{15}$$

As discussed in Lettau and Pelger (2020a), this signal matrix includes additional terms due to the noise, but all conceptual arguments can be made with this simplified signal matrix. A factor can be weak with a small variance. Hence, PCA might not be able to detect it or estimate it precisely. However, if this factor has a large mean, then by choosing γ_{TS} appropriately, the eigenvalues of this factor in the matrix (15) can be pushed up. Thus, the factor can be detected or estimated more precisely.

Similar arguments apply to an α target. The signal matrix would be close (up to the additional correction terms due to noise) to

$$\Sigma_F + \mu_F\mu_F^{\top} + (1+\gamma_{\text{TS}})\,\alpha_{\text{ref}}\alpha_{\text{ref}}^{\top},$$

where α_{ref} is the pricing error of the factors F relative to the reference factors F^{ref}. Hence, we can push up the eigenvalues of weak factors with small variance, that can explain a large part of the pricing errors relative to the reference factors.

The general argument underlying the properties of XS-TS-Target-PCA does not depend on the specific asymptotic regime or rate conditions. In the end, any empirical data set is finite, and the different asymptotic regimes are tools to best describe the finite sample behavior. The fundamental insight is that the eigenvalues of factors can be pushed up by including additional information. This can be theoretically described as the detection of very weak factors, an increase in convergence rates for semi-weak factors or higher efficiency of strong factors. In all cases it results in a better model. The key point is that the finite sample eigenvalues are larger after including the penalty than before. XS-TS-Target-PCA offers a disciplined approach to include additional economic information and to test if this target actually includes information that is not already included in the second moments of the panel of test assets.

3 Empirical results

In this section, we discuss the empirical performance of the cross-sectional time-series target PCA approach using a high-dimensional cross section containing the returns of the $N = 370$ decile portfolios of 37 firm characteristics. To disentangle the value added by the cross-sectional and time-series targets, we first study the performance of the models based on a single target and then the model based on a combination of the two types of targets. Section 3.1 describes the data and the methodology employed to evaluate out-of-sample performance. Section 3.2 considers the approach that exploits a cross-sectional shape target, Section 3.3 the approach that exploits a time-series alpha target, Section 3.4 the approach that combines cross-sectional and time-series targets, and Section 3.5 the approach that exploits a cross-sectional macro target.

3.1 Data and evaluation methodology

We use the same test assets and sample period as Kozak et al. (2020) and Lettau and Pelger (2020b). The dataset contains the monthly returns for the 370 value-weighted decile portfolios obtained from the single sorts of 37 firm-specific characteristics, with decile breakpoints based on NYSE stocks. The sample period spans from November 1963 to December 2017. Table B.1 in Appendix B classifies the 37 characteristics into eight categories.

To construct the cross-sectional macro target, we also use monthly data for the 127 macroeconomic variables from FRED-MD. We construct the innovations to these macroeconomic variables using the following procedure. First, we transform these variables following the method suggested in the appendix of McCracken and Ng (2016) to render them stationary. Next, we standardize the transformed macroeconomic variables such that each of them has zero mean and unit standard deviation. We then apply PCA to these standardized variables and identify *seven* factors that drive their systematic variation.[4] Lastly, we run a first-order vector autoregression on these PCA factors and take the residuals as the seven innovations to the macroeconomic variables. The sample for the macroeconomic innovations spans the same period as the returns of the decile portfolios, from November 1963 to December 2017.

We consider three criteria to evaluate the performance of different latent-factor models. First, the maximum *Sharpe ratio* that can be obtained with a linear combination of the latent factors $\widehat{F}$:

$$\mathrm{SR} = \sqrt{\mu_{\widehat{F}}^{\top} \Sigma_{\widehat{F}}^{-1} \mu_{\widehat{F}}}, \tag{16}$$

where $\mu_{\widehat{F}}$ and $\Sigma_{\widehat{F}}$ are the mean and covariance matrix of the latent factor retirms, respectively. Second, the *root-mean-squared pricing error* across N test assets:

$$\mathrm{RMS}_{\alpha} = \sqrt{\hat{\alpha}^{\top} \hat{\alpha}/N}, \tag{17}$$

[4] We select seven factors using the criterion of Bai and Ng (2002). In particular, we find that seven PCAs of the standardized macroeconomic variables explain 45.4% of their total variation.

where $\hat{\alpha}$ is the vector of OLS intercepts obtained by regressing the test-asset returns on the latent factors. Third, the *average unexplained variance* across N test assets:

$$\overline{\sigma}_e^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{\sigma^2_{\hat{e}_i}}{\sigma^2_{R_i}}, \tag{18}$$

where $\hat{e}_i$ is the residual obtained by regressing the returns of asset i on the latent factors, and $\sigma^2_{\hat{e}_i}$ and $\sigma^2_{R_i}$ are the variance of the residual and return of asset i, respectively.

We evaluate the out-of-sample performance of different factor models using the following rolling-window procedure.[5] For each month t, we first estimate factors returns $\hat{F}$ and factor loadings $\hat{\beta}$ in a rolling window of 20 years ($T = 240$ months) including up to month t. Using these estimated parameters, we compute the out-of-sample factor returns and asset-pricing errors in month $t+1$. We also compute the mean-variance portfolio weights for the factor returns in the rolling window and use them to compute the out-of-sample return of the maximum Sharpe-ratio portfolio for month $t+1$. The mean and variance of the out-of-sample pricing errors are used to calculate the average pricing error and the average magnitude of unexplained idiosyncratic variation, respectively. The Sharpe ratio of the out-of-sample returns of the maximum Sharpe-ratio portfolio is equal to the maximum Sharpe ratio that can be obtained from a linear combination of the latent factors.

In the remainder of this section, we evaluate the empirical performance of our proposed cross-sectional time-series PCA approach. To disentangle the effect of the cross-sectional and time-series targets, we first evaluate the performance of the models based on a single target (cross-sectional or time-series), and then we study the performance of models based on a combination of the two types of targets.

3.2 Cross-sectional shape target

We first consider a *cross-sectional shape target* that exploits the empirical observation that risk premia are monotonic on certain firm-specific characteristics. Therefore, we expect the loadings on asset pricing factors to be monotonic on these characteristics. To incorporate this economic restriction, we consider a cross-sectional target matrix $\Lambda^{\mathrm{XS}} \in \mathbb{R}^{370\times 37}$, whose lth column contains the weights of a long-short portfolio of the lth characteristic, as illustrated in Figure 1. Mathematically, the element on the nth row and lth column of Λ^{XS} is

$$\Lambda^{\mathrm{XS}}_{nl} = \begin{cases} -1 + \frac{2}{9}(i-1) & \text{if asset } n \text{ is the } i\text{th decile portfolio of characteristic } l, \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

[5] The Internet Appendix shows that our findings are robust to evaluating model performance using three-fold cross-validation instead of a rolling-window procedure.

This cross-sectional shape target nudges the latent factors to have weights that fit a monotonic long-short pattern on the decile portfolios of each characteristic. To measure how factor weights fit this monotonic pattern, we define the cross-sectional metric of the ith factor as

$$\text{XS-Metric}(\hat{\omega}_i) = 100 \times \frac{\|P_\Lambda^{\text{XS}} \hat{\omega}_i\|_2}{\|\hat{\omega}_i\|_2}, \tag{20}$$

where $\hat{\omega}_i$ is the estimated factor weight vector and P_Λ^{XS} is the cross-sectional projection matrix associated with the target Λ^{XS}. Therefore, the closer the cross-sectional metric of a latent factor is to 100, the better it is spanned by the economically motivated factors.

3.2.1 Number of factors and penalty parameter

Figure 2 illustrates how the out-of-sample performance of the cross-sectional shape PCA (XS-shape-PCA) model varies with the number of factors and the penalty. In particular, Panel A gives a line plot depicting the Sharpe ratio of each model on the vertical axis as a function of the penalty, γ_{XS}, on the horizontal axis. Panels B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across 37 top-minus-bottom portfolios, and unexplained variance across the 370 decile portfolios, respectively. For each heatmap, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, γ_{XS}. Herein, root-mean-squared pricing errors are reported in basis points and unexplained variances in percentage.

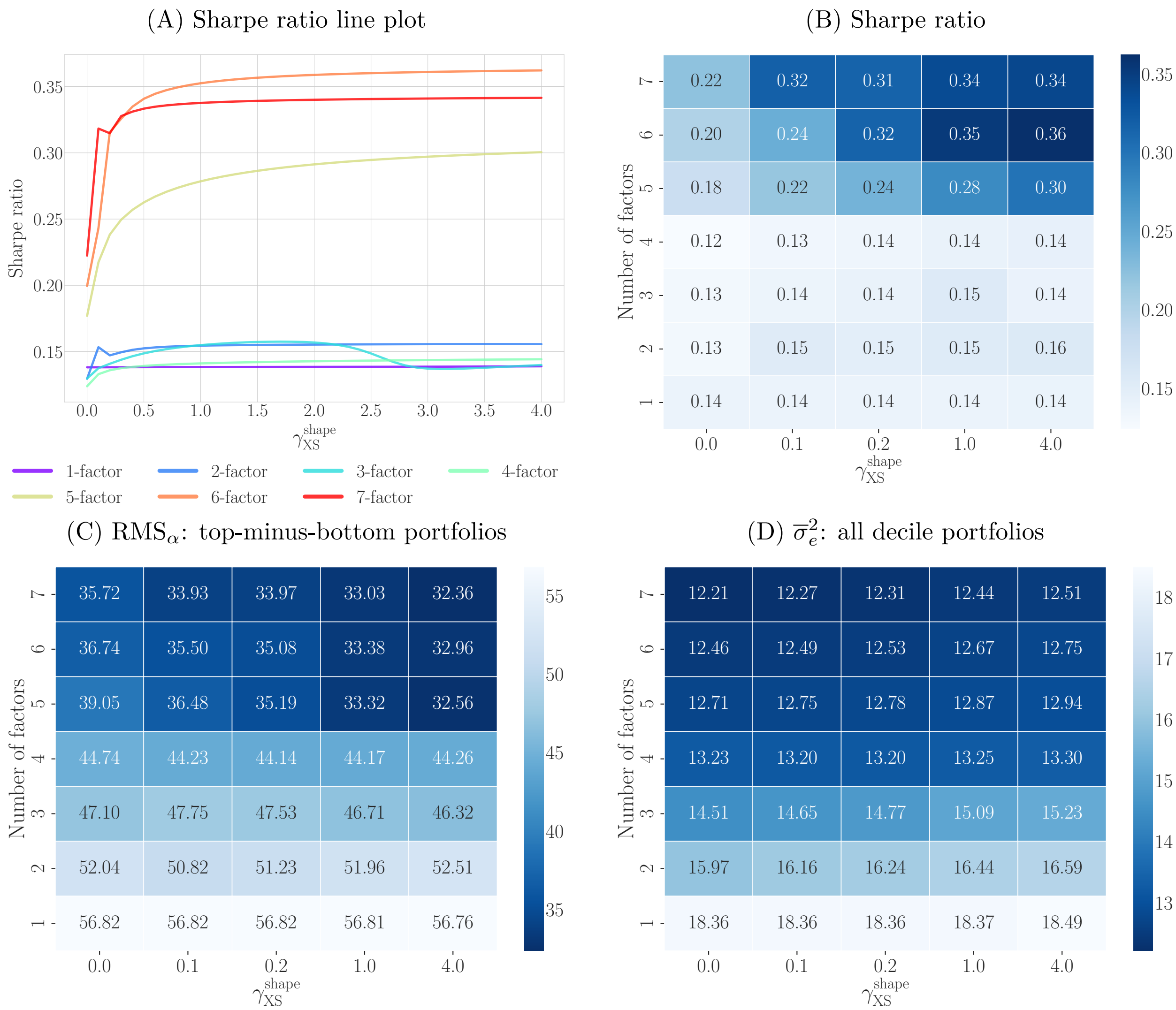
Panels A and B in Figure 2 show that the shape target helps to substantially increase the Sharpe ratio of high-dimensional models with five or more factors. For example, the Sharpe ratio of the five-factor model increases from 0.18 to 0.28 when the penalty γ_{XS} increases from zero (conventional PCA) to one (XS-shape-PCA). This is a substantial 57.4% increase in the Sharpe ratio, which suggests that the fifth factor identified by the cross-sectional target PCA model may be a weak factor that is important for asset pricing, but missed by conventional PCA.

Panel C reports the heatmap for the root-mean-squared pricing error of the 37 top-minus-bottom portfolios. Each of these 37 portfolios goes long the top decile and short the bottom decile of a characteristic, and thus, they can be interpreted as the target-managed portfolios. The results in Panel C are consistent with those for the Sharpe ratio in Panels A and B. In particular, the average pricing error of a low-dimensional model with less than five factors does not vary substantially with the penalty parameter γ_{XS}, while those of the high-dimensional five-, six-, and seven-factor models decrease monotonically with γ_{XS}. For example, the root-mean-squared pricing error across the top-minus-bottom portfolios is 39.05 for the five-factor PCA model, but it is only 33.32 (a 14.7% reduction) for the XS-shape-PCA with $\gamma_{\text{XS}} = 1$.

Panel D in Figure 2 shows that PCA delivers the lowest average unexplained variance across all 370 decile portfolios, which is not surprising because PCA is designed to explain the time-series variation of the portfolio returns. More importantly, the average unexplained variance does not increase substantially with the penalty parameter γ_{XS}. For example, the average unexplained

Figure 2: Performance of cross-sectional shape PCA models

This figure illustrates the out-of-sample performance of cross-sectional shape PCA (XS-shape-PCA) models with up to seven factors and for different values of the penalty parameter. Panel A gives a line plot depicting the Sharpe ratio of each model on the vertical axis as a function of the penalty parameter, γ_{XS}, on the horizontal axis. Panels B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, and average unexplained variance across the 370 decile portfolios, respectively. For each heatmap, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, γ_{XS}. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



variance is 12.71% for the five-factor PCA model; when γ_{XS} increases to one, it only increases by 1.3% to 12.87%. Therefore, even if the shape target nudges the PCA factors to span different subspaces of asset returns, the XS-shape-PCA factors explain almost the same amount of time-series variation as the PCA factors.

Overall, we find that the shape target generally helps to improve the performance of the conventional PCA model. In particular, performance across all three criteria typically improves

Table 1: Performance of cross-sectional shape PCA and other models

This table summarizes the performance of three- and five-factor models including the Fama-French, PCA, and XS-shape-PCA models. The benchmark penalty parameter on the cross-sectional shape target is $\gamma_{\text{XS}} = 1$. Panels A and B report the in-sample and out-of-sample results, respectively. For each model, the five columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the top-minus-bottom portfolios, unexplained variance across all decile portfolios, and the cross-sectional metric of the stochastic discount factor, respectively. Root-mean-squared pricing errors are reported in basis points and unexplained variances in percentage. Bold numbers indicate the best-performing models.

(A) In-sample

Model	Sharpe ratio	RMS_α for TMB portfolios	$\overline{\sigma}_e^2$ for all decile portfolios	XS metric
Three-factor models				
FF	0.21	57.33	14.27	-
PCA	0.17	55.78	**12.57**	66.55
XS-shape-PCA	**0.23**	**45.95**	13.27	**98.02**
Five-factor models				
FF	0.32	49.45	13.27	-
PCA	0.25	42.89	**10.80**	80.99
XS-shape-PCA	**0.33**	**38.35**	10.91	**97.87**

(B) Out-of-sample

Model	Sharpe ratio	RMS_α for TMB portfolios	$\overline{\sigma}_e^2$ for all decile portfolios	XS metric
Three-factor models				
FF	**0.16**	52.05	16.63	-
PCA	0.13	47.10	**14.51**	72.45
XS-shape-PCA	0.15	**46.71**	15.09	**96.57**
Five-factor models				
FF	**0.31**	36.92	15.64	-
PCA	0.18	39.05	**12.71**	77.97
XS-shape-PCA	0.28	**33.32**	12.87	**96.69**

when increasing the penalty parameter, compared to the conventional PCA case with $\gamma_{\text{XS}} = 0$. This suggests that the shape target is informative for cross-sectional expected returns.

Finally, Figure 2 shows that a five-factor model offers a good tradeoff between out-of-sample performance and parsimony. For instance, Panel C shows that there is a great reduction in root-mean-squared pricing error when increasing the number of factors from four to five, but there is no reduction when increasing the number of factors to six. Figure 2 also shows that the performance of the XS-shape-PCA models stabilizes when γ_{XS} reaches one. Therefore, in the next section we use the case with five factors and $\gamma_{\text{XS}} = 1$ as a benchmark for comparison with other models.

3.2.2 Out-of-sample comparison

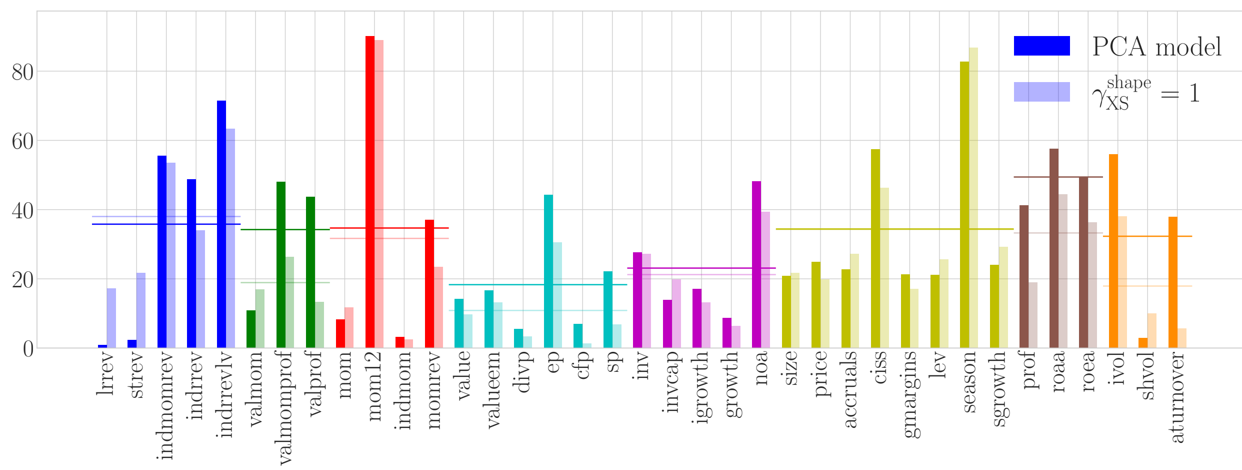
Table 1 compares the performance of the XS-shape-PCA models with that of the Fama-French and conventional PCA models. Panels A and B contain the in-sample and out-of-sample results, respectively. In each panel, the five columns report the model acronym, Sharpe ratio, root-mean-squared pricing error across the top-minus-bottom portfolios, average unexplained variance across all decile portfolios, and cross-sectional shape metric of each model, respectively.

Table 1 shows that the shape target helps to improve the pricing performance of PCA models. In particular, the XS-shape-PCA model has higher Sharpe ratio and lower pricing error than conventional PCA. This improvement is highly consistent, holding both in-sample and out-of-sample as well as for three- and five-factor models. However, the improvement is particularly pronounced for the case with five factors, which again suggests that the XS-shape-PCA identifies a weak factor among the top five that while being important for asset pricing, is missed by conventional PCA.

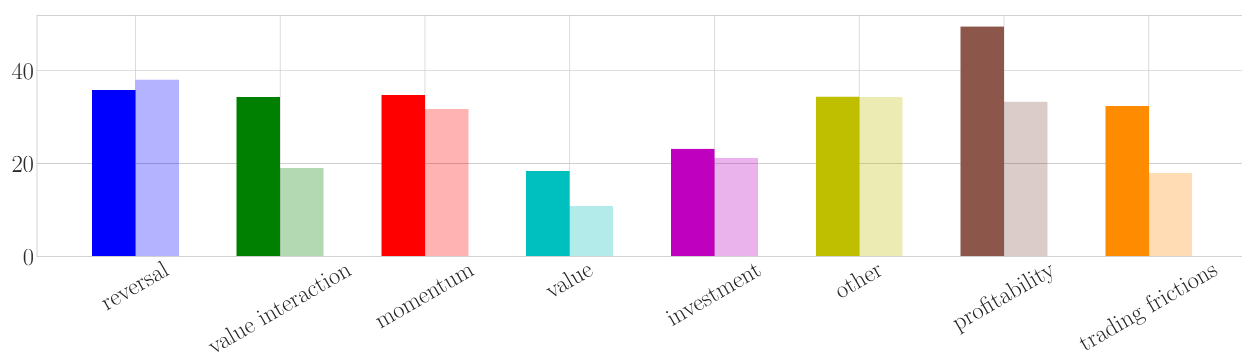
Figure 3: Pricing errors for top-minus-bottom portfolios, five-factor PCA and XS-shape-PCA

This figure illustrates the out-of-sample pricing errors (in absolute value) for the 37 top-minus-bottom portfolios of the five-factor PCA model and XS-shape-PCA model with $\gamma_{\text{XS}} = 1$. Panel A shows the pricing error of each top-minus-bottom portfolio, and the horizontal axis depicts the 37 characteristics. Panel B shows the average pricing error for portfolios in the same category, and the horizontal axis depicts the eight categories of characteristics. The dark and light bars correspond to the PCA and XS-shape-PCA models, respectively, and the dark and the light horizontal lines in Panel A show the average pricing errors for all portfolios in the corresponding categories of the PCA and XS-shape-PCA models, respectively. Pricing errors are reported in basis points.

(A) Pricing error of each portfolio



(B) Average pricing error of portfolios in each category

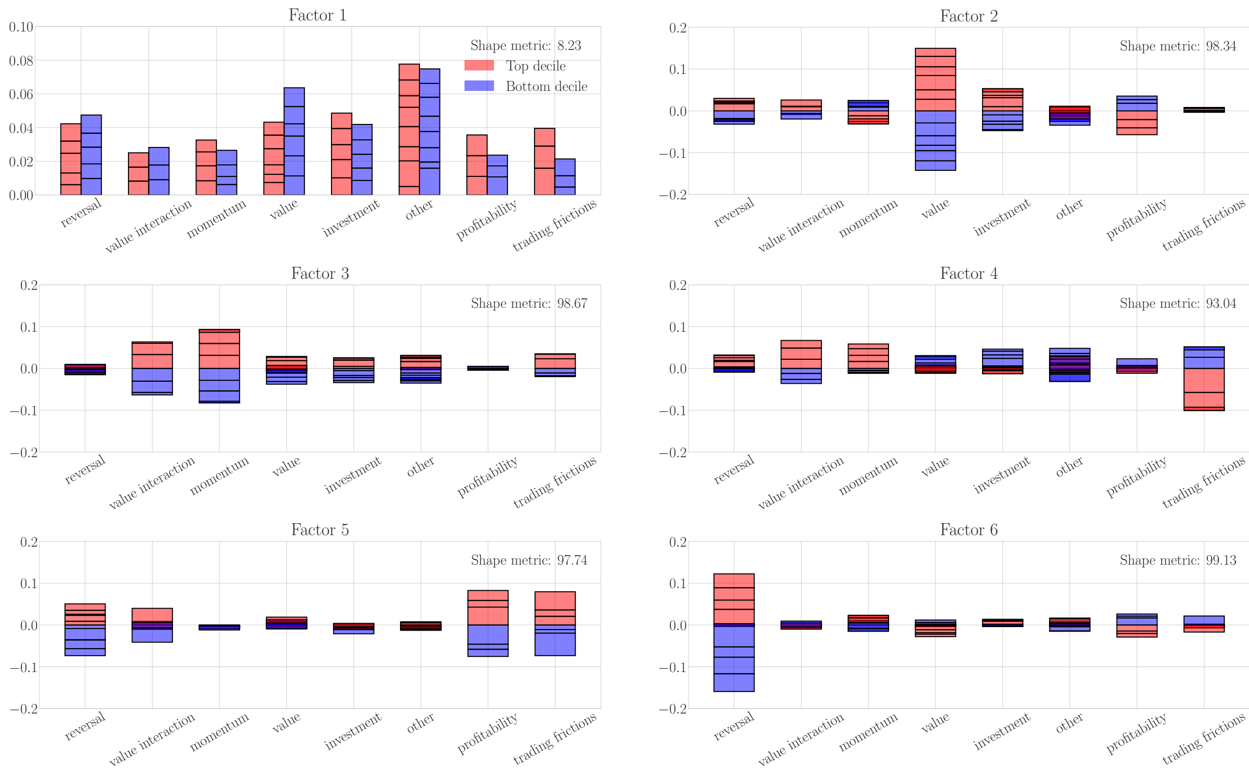


Note also that the enhanced pricing performance of XS-shape-PCA models can be traced back to their higher cross-sectional metrics. For example, the out-of-sample cross-sectional metric of the five-factor XS-shape-PCA model is 96.69, which is higher than that of the conventional PCA model, 77.97. This implies that the weights of the XS-shape-PCA factors have a stronger long-short pattern in characteristics than those of the PCA factors. Therefore, it is not surprising that the XS-shape-PCA models have similar Sharpe ratios as the Fama-French models, whose factors are also constructed by exploiting various long-short restrictions. In particular, the out-of-sample Sharpe ratio of the five-factor XS-shape-PCA model, 0.28, is very close to that of the FF5 model, 0.31.

Figure 3 compares the pricing errors (in absolute value) for all top-minus-bottom portfolios of the five-factor PCA model and the XS-shape-PCA model with $\gamma_{\text{XS}} = 1$. Panel A shows the

Figure 4: XS-shape-PCA factor weights on extreme-decile portfolios

This figure shows the in-sample weights of the first six XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$ on the extreme-decile portfolios, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive average returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.



pricing error for each portfolio and Panel B shows the average pricing error for portfolios in each category. The dark and light bars correspond to the PCA and XS-shape-PCA models, respectively. This figure shows that the XS-shape-PCA model delivers *uniformly* lower pricing errors for top-minus-bottom portfolios in all categories except reversal. In particular, the shape target results in over 40% lower average pricing error in three out of the eight characteristic categories.

Since the XS-shape-PCA models explain the average returns of the long-short portfolios of characteristics, which are known in the literature to earn high risk premia, it is expected that the XS-shape-PCA factors pick up these directional risk premia and deliver a high Sharpe ratio. Moreover, PCA failing to fully identify the directional risk premia implies that some of these premia are carried by weak factors. The cross-sectional shape target explicitly increases the signal strength of these factors, and thus they appear as the higher-order XS-shape-PCA factors. Therefore, the performance of low-dimensional PCA and XS-shape-PCA models is similar, while the high-dimensional XS-shape-PCA models have substantially higher Sharpe ratios than their PCA counterparts.

3.2.3 Structure and composition of factors

We now study the composition of the XS-shape-PCA factors. Figure 4 graphs the in-sample weights on the extreme-decile portfolios of the first six XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$. The vertical axis depicts the weights and the horizontal axis gives the characteristic categories.[6] All factors are normalized to have positive average returns. Weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

Figure 4 shows that the shape target nudges the factors to have weights that are monotonic on the characteristics. In particular, the figure shows that the weights of every XS-shape-PCA factor (except the first) display a strong long-short pattern on the decile portfolios of characteristics. Indeed, the average cross-sectional metric of XS-shape-PCA factors two to six is 97.38, while that of PCA factors two to six is 73.16, which is much lower.[7] This demonstrates that XS-shape-PCA successfully incorporates the economic prior behind the shape constraints into the construction of latent factors.

The first XS-shape-PCA factor resembles an equally-weighted market portfolio. In fact, both the first PCA and XS-shape-PCA factors have correlations with the market factor above 0.99. The second factor can be labeled as a value factor because its weights have a strong long-short pattern on the value extreme portfolios. The third factor assigns large long-short weights to portfolios in the momentum and value interaction categories. The weights on the top- and bottom-decile portfolios of the fourth factor are slightly less symmetric, and for some categories, such as others, the long-short pattern is weaker. The fifth factor has large long-short weights on extreme portfolios in the profitability, trading friction, reversal, and value interaction categories, and the sixth factor can be labeled as a reversal factor because it is almost exclusively composed of long-short reversal portfolios.

We find that the second and third XS-shape-PCA factors resemble the third and fourth PCA factors, whose cross-sectional metrics are 82.16 and 89.92, respectively. Since these PCA factors have high cross-sectional metrics, and thus, approximately satisfy the restriction of the shape target, they also appear as XS-shape-PCA factors, with their signals increased and the long-short pattern strengthened so that they become lower-order XS-shape-PCA factors.

However, the XS-shape-PCA factors and PCA factors are different in general. For example, Panels A and B in Figure 5 give heatmaps for the weights on all decile portfolios of the fifth PCA and XS-shape-PCA factors, respectively.[8] This figure shows that the weights of the XS-shape-PCA factor have a stronger monotonic pattern in the characteristics than the PCA factor. Moreover, this pattern is economically significant because the monthly average return and Sharpe ratio of the

[6]For comparison, Panel A of Figure D.5 in Appendix D graphs the in-sample weights on the extreme portfolios of the first six PCA factors. In addition, Panels A and B of Figure **??** in the Internet Appendix illustrate the in-sample weights of PCA and XS-shape-PCA factors on the extreme-decile portfolios of each characteristic, respectively.

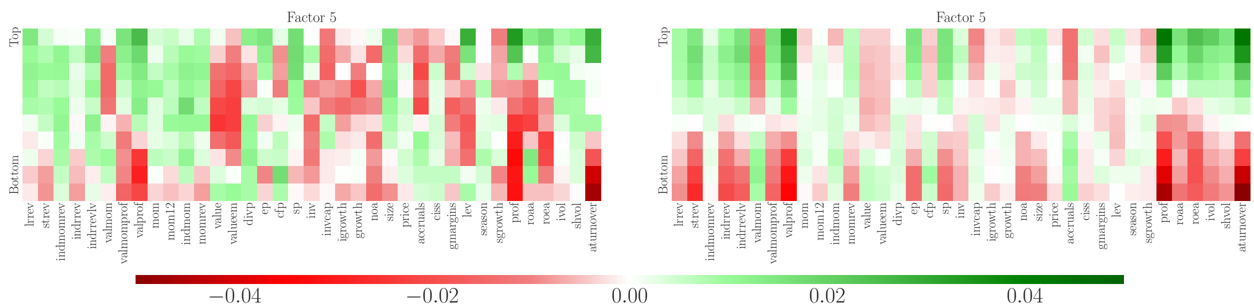
[7]Panel A of Figure D.5 in Appendix D shows the cross-sectional metrics of the PCA factors.

[8]For completeness, Panels A and B of Figure D.2 in Appendix D give heatmaps for the weights on all decile portfolios of the first six PCA and XS-shape-PCA factors, respectively.

Figure 5: Weights on all decile portfolios of the fifth factor

Panels A and B give heatmaps for the in-sample weights on all decile portfolios of the fifth PCA factor and the fifth XS-shape-PCA factor with $\gamma_{\mathrm{XS}} = 1$, respectively. The horizontal axis depicts all characteristics and the vertical axis the bottom- to top-decile portfolios. Positive (negative) weights are shown in green (red).

(A) PCA (B) XS-PCA



fifth XS-shape-PCA factor, which are 0.27% and 0.22, respectively, are higher than those of the fifth PCA factor, which are 0.17% and 0.12, respectively.

To summarize this section, we find that the cross-sectional shape PCA models, especially the high-dimensional ones, have superior cross-sectional pricing performance compared to their PCA counterparts, while the two sets of models explain a similar amount of time-series variation of the decile portfolios. Specifically, the high-dimensional five-, six-, and seven-factor XS-shape-PCA models deliver higher Sharpe ratios and lower root-mean-squared pricing errors across the top-minus-bottom portfolios than their PCA counterparts. Importantly, the XS-shape-PCA models outperform the PCA models because the cross-sectional shape target nudges the latent factors to explain the expected returns of the long-short portfolios of characteristics, which carry the directional risk premia.

3.3 Time-series alpha target

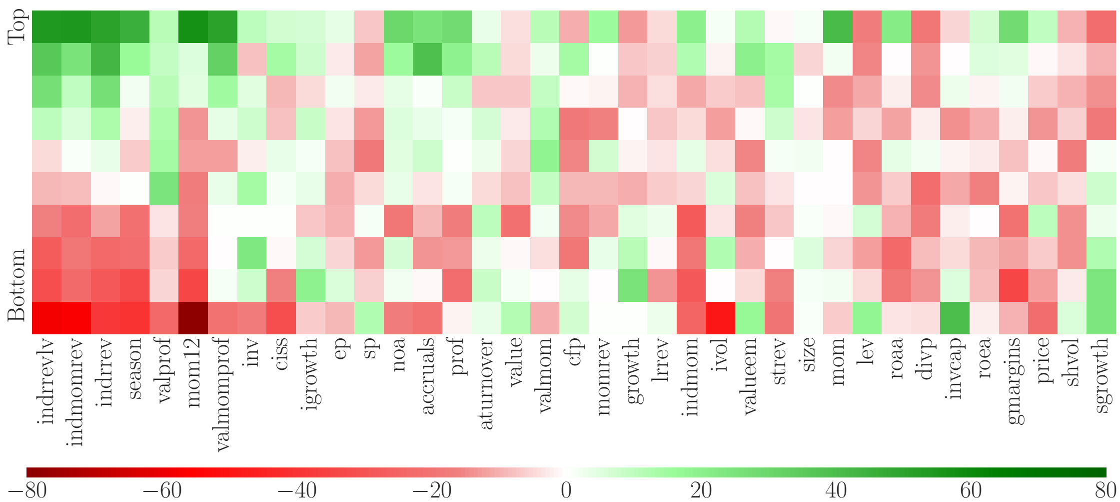
We now consider a *time-series alpha target*, which nudges the latent factors to explain risk premia that a candidate asset-pricing model fails to explain. We consider the five-factor Fama-French model (FF5) as the candidate model in our empirical analysis. Thus, we define the target time-series matrix G^{TS} as in Equation (10) after replacing F^{ref} with the returns of the five Fama-French factors.[9]

Figure 6 graphs the in-sample alphas of all decile portfolios with respect to FF5. The horizontal axis gives the characteristics in descending order of Sharpe ratio of top-minus-bottom portfolios. The figure shows that the alphas with respect to FF5 of the decile portfolios, especially those of characteristics with high Sharpe-ratio top-minus-bottom portfolios, display a strong monotonic pattern, which suggests that the FF5 model cannot fully explain the cross-section of the 370 decile portfolios. Therefore, there is room for time-series alpha PCA to identify factors that can

[9]Lettau and Pelger (2020b) evaluate the RP-PCA models, obtained using the time-series mean target. For completeness, we also report results for the RP-PCA models in the Internet Appendix.

Figure 6: Alphas of decile portfolios with respect to FF5

This figure illustrates the in-sample alphas of the 370 decile portfolios with respect to FF5 in basis points. The horizontal axis gives the characteristics in descending order of Sharpe ratio of top-minus-bottom portfolios. The vertical axis depicts the bottom- to top-decile portfolios. Positive (negative) alphas are shown in green (red).



explain the risk premia in the 370 decile portfolios that is unexplained by the five Fama-French factors.

To evaluate whether the economic restriction of the time-series target is incorporated into the latent factors, we define the time-series metric of the K latent factors as

$$\text{TS-Metric} = 100 \times \frac{\|P_G P_{\hat{F}} R\|_F}{\|P_G R\|_F}, \tag{21}$$

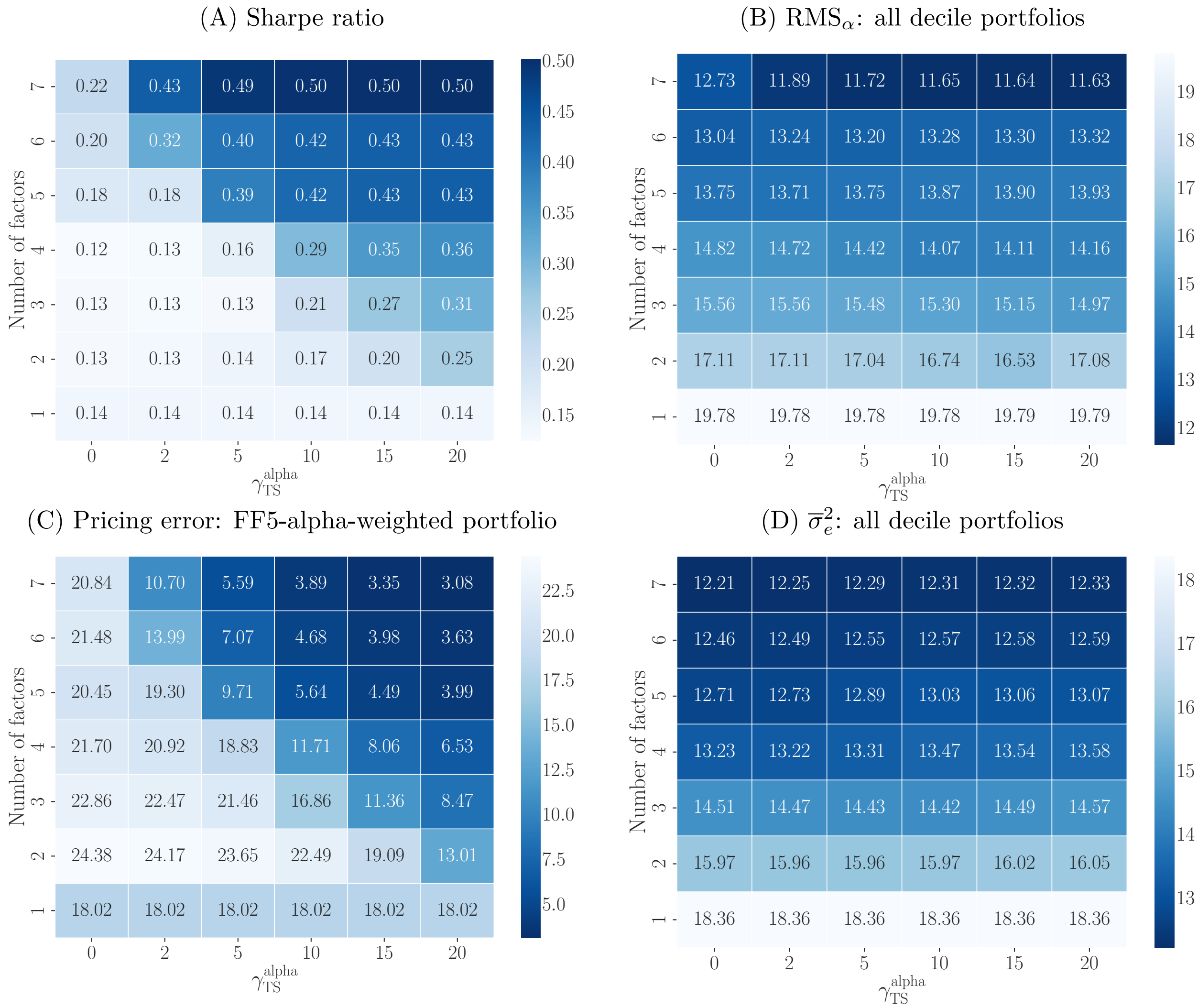
where $P_{\hat{F}} = \hat{F}(\hat{F}^\top \hat{F})^{-1}\hat{F}^\top$ is the matrix that projects N-dimensional vectors onto the column space of $\hat{F}$. The time-series metric gauges how well the factors incorporate the information in the time-series target. In particular, if the time-series metric of the latent factors is close to 100, then the latent factors successfully capture risk premia information identified by the time-series projection matrix P_G.

3.3.1 Number of factors and penalty parameter

Figure 7 illustrates how the out-of-sample performance of the time-series alpha PCA (TS-alpha-PCA) model varies with the number of factors and the penalty parameter. Panel A gives the heatmap for the out-of-sample Sharpe ratio and Panel B for the out-of-sample root-mean-squared pricing error across the 370 decile portfolios. Note that the time-series alpha PCA is designed to explain risk premia that are not explained by the five Fama-French factors. To measure how successful the approach is in achieving this goal, Panel C gives the heatmap for the out-of-sample pricing error for the *FF5-alpha-weighted portfolio*, which is a portfolio whose weight on each decile

Figure 7: Performance of time-series alpha PCA models

This figure illustrates how the out-of-sample performance of the time-series alpha PCA model varies with the number of factors and the penalty parameter. Panel A gives the heatmap for the Sharpe ratio and Panel B for the root-mean-squared pricing error across the 370 decile portfolios. Panel C gives the heatmap for the pricing error for the FF5-alpha-weighted portfolio, which is a portfolio whose weight on each decile is proportional to the decile's alpha with respect to FF5. Finally, Panel D gives the heatmap for the average unexplained variance across the 370 decile portfolios. In each panel, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, $\gamma_{\text{TS}}^{\text{alpha}}$. Pricing errors are reported in basis points and average unexplained variances in percentage.



is proportional to the decile's alpha with respect to FF5.[10] Finally, Panel D gives the heatmap for

[10]To construct the FF5-alpha-weighted portfolio, we first regress the returns of all decile portfolios on FF5 and obtain their alphas. We then construct the FF5-alpha-weighted portfolio by going long decile portfolios with positive alphas and going short those with negative alphas. The portfolio weights on the long and short legs are proportional to the alphas and are standardized such that the total weight of each leg is one. The FF5-alpha-weighted portfolio has an in-sample average return of 35.9 basis points and an out-of-sample average return of 19.5 basis points.

unexplained variance across the 370 decile portfolios. In each panel, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, $\gamma_{\text{TS}}^{\text{alpha}}$.[11]

Panel A in Figure 7 shows that the time-series alpha target substantially increases the Sharpe ratios of latent-factor models. For example, the Sharpe ratio of the five-factor model more than doubles from 0.18 to 0.43 when we increase the penalty parameter $\gamma_{\text{TS}}^{\text{alpha}}$ from zero (conventional PCA) to 15.

Given that the time-series alpha target helps to increase the out-of-sample Sharpe ratio, one may expect that it will also help to reduce the root-mean-squared pricing error. However, Panel B in Figure 7 shows that the time-series alpha target does not substantially affect the root-mean-squared pricing error for the 370 decile portfolios. For instance, for the five-factor model, when $\gamma_{\text{TS}}^{\text{alpha}}$ increases from zero to 15, the pricing error experiences only a minor increase, from 13.75 to 13.90. The explanation for this is that the objective of the time-series alpha PCA is *not* to explain the expected returns of the decile portfolios, but rather to explain the *alpha* of the decile portfolios with respect to the FF5 model.

Indeed, Panel C in Figure 7 shows that the time-series alpha target helps to substantially decrease the pricing error for the FF5-alpha-weighted portfolio. Specifically, PCA models, even the high-dimensional ones, fail to explain the expected return of the FF5-alpha-weighted portfolio. For example, the pricing errors of the three- and five-factor PCA models are 22.86 and 20.45. However, when $\gamma_{\text{TS}}^{\text{alpha}}$ increases to 15, the pricing errors of the three- and five-factor models are just 11.36 and 4.49, which are 50.3% and 78.0% lower than their PCA counterparts, respectively.

Panel D in Figure 7 shows that the PCA model delivers the lowest average unexplained variance for almost every number of factors. However, similar to the cross-sectional shape target, the time-series alpha target also has a limited impact on the time-series explanatory power of models. For example, the average unexplained variance is 12.71% for the five-factor PCA model; when $\gamma_{\text{TS}}^{\text{alpha}}$ increases to 15, it only increases by 2.8% to 13.06%.

Overall, the time-series alpha target helps to substantially increase the out-of-sample Sharpe ratio of latent-factor models and although it does not help to reduce the root-mean-squared pricing error of the 370 decile portfolios, it does substantially reduce the pricing error of the FF5-alpha-weighted portfolio, without reducing the time-series explanatory power of the models.

Finally, for parsimony we focus the rest of our discussion on the five-factor model. Also, Figure 7 shows that the performance of the TS-alpha-PCA models stabilizes when $\gamma_{\text{TS}}^{\text{alpha}}$ reaches 15. In addition, Figure D.8 in Appendix D illustrates that the performance of the TS-mean-PCA models also stabilizes when $\gamma_{\text{TS}}^{\text{mean}}$ reaches 15. Therefore, in the next section we use the case with five factors and $\gamma_{\text{TS}} = 15$ as a benchmark for comparison with other models.

[11]Figure D.8 in Appendix D is the counterpart of Figure 7 to illustrate how the out-of-sample performance of the time-series *mean* PCA model, which is equivalent to the RP-PCA model of Lettau and Pelger (2020b), varies with the number of factors and the penalty parameter.

Table 2: Performance of time-series alpha PCA and other models

This table summarizes the performance of three- and five-factor models including the Fama-French, PCA, TS-mean-PCA, and TS-alpha-PCA models. The benchmark penalty parameter on the time-series target is $\gamma_{\text{TS}} = 15$. Panels A and B report the in-sample and out-of-sample results, respectively. For each model, the six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, pricing error for the FF5-alpha-weighted portfolio, average unexplained variances across the 370 decile portfolios, and time-series alpha metric, respectively. Pricing errors are reported in basis points and unexplained variances in percentage. Bold numbers indicate the best-performing models.

Models	Sharpe ratio	RMS_α for all decile portfolios	Pricing error for FF5-alpha-weight. portfolio	$\overline{\sigma}_e^2$ for all decile portfolios	TS-alpha metric
		Panel A: in-sample			
Three-factor models					
FF	0.21	17.84	43.00	14.27	-
PCA	0.17	16.94	42.28	**12.57**	38.71
TS-mean-PCA	**0.26**	**16.73**	46.19	12.68	14.92
TS-alpha-PCA	0.16	18.28	**32.98**	12.62	**74.31**
Five-factor models					
FF	0.32	16.09	46.22	13.27	-
PCA	0.25	14.23	34.01	**10.80**	49.14
TS-mean-PCA	0.57	13.44	15.66	10.95	88.37
TS-alpha-PCA	**0.66**	**12.40**	**4.42**	11.07	**99.35**
		Panel B: out-of-sample			
Three-factor models					
FF	0.16	17.23	20.86	16.63	-
PCA	0.13	15.56	22.86	14.51	39.60
TS-mean-PCA	0.20	**14.13**	17.92	14.56	65.14
TS-alpha-PCA	**0.27**	15.15	**11.36**	**14.49**	**87.63**
Five-factor models					
FF	0.31	13.66	28.41	15.64	-
PCA	0.18	13.75	20.45	**12.71**	45.29
TS-mean-PCA	**0.46**	**11.84**	10.23	12.98	86.22
TS-alpha-PCA	0.43	13.90	**4.49**	13.06	**99.37**

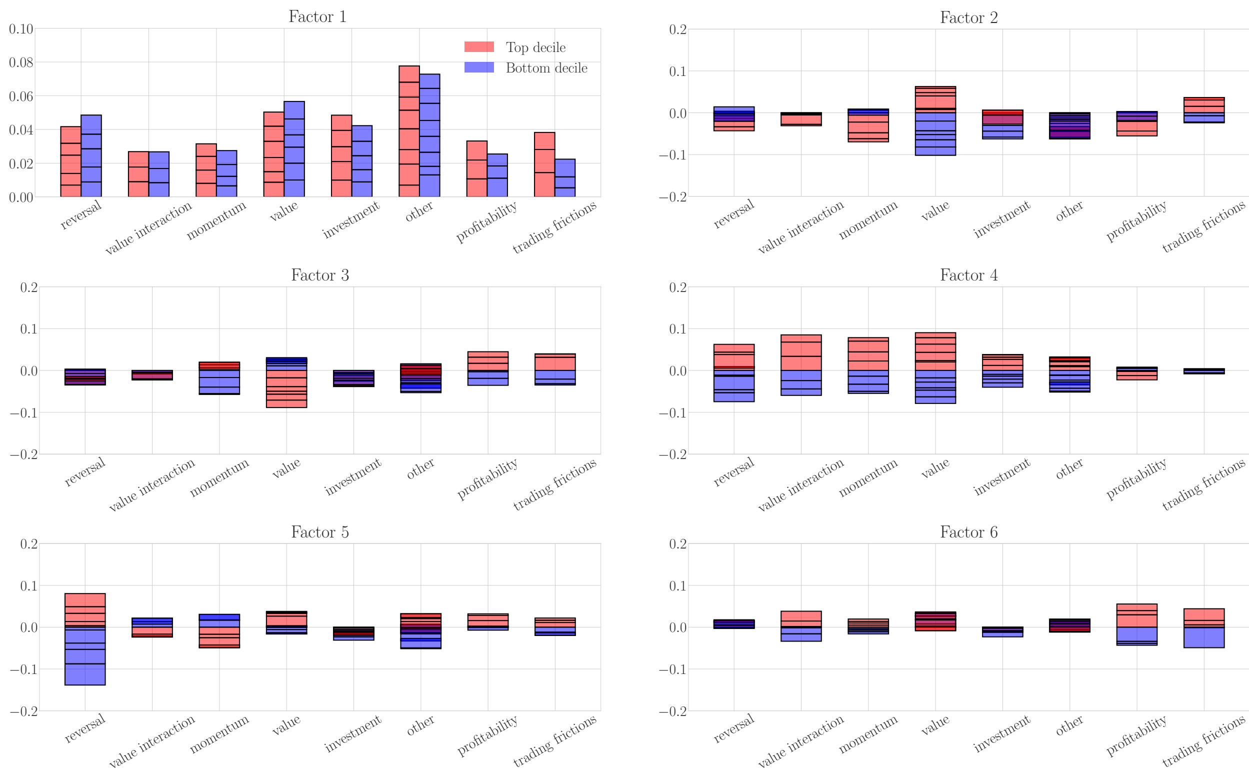
3.3.2 Out-of-sample comparison

Table 2 compares the performance of the TS-alpha-PCA models with that of other three- and five-factor models including the Fama-French, PCA, and time-series mean PCA (TS-mean-PCA) models. Panels A and B give the in-sample and out-of-sample results, respectively. In each panel, the six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, pricing error for the FF5-alpha-weighted portfolio, average unexplained variance across the 370 decile portfolios, and time-series alpha metrics of each model, respectively.

Table 2 shows that each of the time-series targets helps to improve performance with respect to the particular criterion they are designed to address. For instance, the time-series *mean* target helps to reduce the root-mean-squared error of the 370 decile portfolios because the time-series mean target is designed to explain the risk premia of these portfolios. However, the time-series *alpha* target helps to reduce the pricing error for the FF5-alpha-weighted portfolio, which it is designed to explain. Note also that the TS-alpha-PCA model has the highest TS-alpha-metric,

Figure 8: TS-alpha-PCA factor loadings on extreme-decile portfolios

This figure shows the loadings of the first six TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$ on the extreme-decile portfolios, with the horizontal axis depicting the categories of the characteristics. All factors are normalized to have positive average returns. The loadings on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total loading of a category, and the black lines indicate the contribution of each portfolio in the category.



which measures how well a model incorporates risk premia that are not explained by the FF5 model. The two time-series targets (mean and alpha) help to explain different risk premia, but both help the latent factors to achieve higher Sharpe ratios as they explicitly include risk premia information. Therefore, it is not surprising that the two time-series target PCA models outperform the Fama-French and PCA models, neither of which explicitly exploits any risk premia information.

3.3.3 Structure and composition of factors

We now study the composition of the TS-alpha-PCA factors. Figure 8 graphs the in-sample weights on the extreme-decile portfolios of the first six TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. All factors are normalized to have positive average returns.[12] Weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

[12]To facilitate comparison, Panels A and B of Figure D.3 in Appendix D graph the in-sample weights on the extreme-decile portfolios of the first six PCA and TS-alpha-PCA factors. In addition, Panel B of Figure D.4 in Appendix D graphs the in-sample weights on the extreme-decile portfolios of the first six benchmark TS-*mean*-PCA factors with $\gamma_{\text{TS}}^{\text{mean}} = 15$.

Table 3: Regression results of pure-alpha portfolio on TS-alpha-PCA and PCA factors

This table reports the results of regressing the returns of the pure-alpha portfolio on the first five TS-alpha-PCA factors and the first five PCA factors, respectively. The first and second columns report the model acronyms and measures of interest, respectively. The following six columns report the parameters corresponding to the intercept and the factors, and the last column reports the regression R-squared. For each model, the first row reports the regression coefficients, the second row reports the expected return of the pure-alpha portfolio explained by each factor, and the last row reports the R-squared when each factor is excluded from the regression. The intercept and expected return explained by each factor are reported in basis points.

Model	Measure	Intercept	Factors					R^2 (%)
			1	2	3	4	5	
Alpha-TS-PCA	Coefficient	7.75	0.00	−0.33	0.24	0.47	0.33	88.20
	Return explained		0.09	−1.27	4.81	25.39	9.43	
	R^2 when excluded (%)		88.20	36.07	53.14	30.56	74.20	
PCA	Coefficient	38.54	−0.00	−0.05	−0.09	0.28	0.10	56.91
	Return explained		−0.11	−1.52	−0.50	8.07	1.75	
	R^2 when excluded (%)		56.90	54.00	49.79	12.99	54.15	

Comparing the composition of the TS-alpha-PCA factors to that of the conventional PCA factors depicted in Panel A of Figure D.3 in Appendix D, we observe that the fourth and fifth TS-alpha-PCA factors display a long-short pattern on the characteristics in the reversal category that is missing in the conventional PCA factors. This is not surprising as the TS-alpha-PCA factors are designed to explain risk premia that are not explained by the FF5 model such as those associated with the reversal characteristics.

Comparing the weights of the TS-alpha-PCA factors to those of the TS-mean-PCA factors depicted in Panel B of Figure D.4 in Appendix D, we find that the weights of the TS-alpha-PCA factors on decile portfolios in the value, investment, and profitability categories tend to be smaller than those of the TS-mean-PCA factors. This is not surprising because the FF5 factor explains well the decile portfolios in these categories and the TS-alpha-PCA factors are designed to explain risk premia *not* captured by the FF5 factor model. However, we find that the TS-alpha-PCA factors have large long-short weights on decile portfolios in categories that are not explained by the FF5 model. For instance, the weights of the fourth TS-alpha-PCA factor have a strong long-short pattern in the decile portfolios in the reversal, value interaction, and momentum categories. Similarly, the fifth TS-alpha-PCA factor almost only loads on reversal portfolios, which, again, are missed by FF5.[13]

Finally, comparing the composition of the TS-alpha-PCA factors to that of the XS-shape-PCA factors depicted in Figure 4, we observe that the weights on the extreme decile portfolios of the TS-alpha-PCA factors are less monotonic in the characteristics than those of the XS-shape-PCA factor, which are designed to capture precisely such monotonicity. For instance, the weights of the third TS-alpha-PCA factor are not monotonic in the investment decile portfolios.

[13]Recall that the sixth XS-shape-PCA factor is also a reversal factor. Unlike the cross-sectional shape target, the time-series alpha target explicitly increases the signal strength of factors with high risk premia, which justifies why the reversal factor is a lower-order TS-alpha-PCA factor.

Given that the fourth and fifth TS-alpha-PCA factors load heavily on the characteristics in the reversal category, we expect that they will capture risk premia missed by FF5. To verify this, we first construct a pure-alpha portfolio, which only earns risk premia missed by FF5, by regressing the returns of the FF5-alpha-weighted portfolio on FF5 and using the residuals of this regression as the returns of the pure-alpha portfolio. The in-sample average monthly return of this portfolio is 46.22 basis points. We then regress the returns of this portfolio on the first five TS-alpha-PCA factors. For comparison, we also regress the returns of the pure-alpha portfolio on the first five PCA factors. Table 3 reports the results of these two regressions. For each model, the first row reports the regression coefficients, the second row reports the expected return of the pure-alpha portfolio explained by each factor, and the last row reports the R-squared when each factor is excluded from the regression. The intercept and expected return explained by each factor are reported in basis points.

Consistent with the previous results, Table 3 demonstrates that compared to the PCA model, the TS-alpha-PCA model better explains the expected return of the pure-alpha portfolio. Specifically, the pricing error for the pure-alpha portfolio of the TS-alpha-PCA models is only 7.8 basis points, while that of the PCA model is 38.5 basis points. Most importantly, the fourth and fifth TS-alpha-PCA factors explain 25.39 and 9.43 basis points, thus explaining over 75% of the expected return of the pure-alpha portfolio. Therefore, the fourth and fifth TS-alpha-PCA factors, which exploit long-short restrictions on characteristics not exploited by FF5, successfully capture risk premia missed by FF5.[14] In addition, we verify that the second TS-alpha-PCA factor is a time-series factor and only explains the time-series variation of the pure-alpha portfolio. In particular, the R-squared decreases substantially from 88.20% to 36.07% when excluding it from the regression.

To summarize this section, we find that the time-series alpha PCA models have superior cross-sectional pricing ability than their PCA counterparts, while they explain a similar amount of time-series variation of the decile portfolios. Specifically, the Sharpe ratios of the TS-alpha-PCA models are substantially higher than those of the PCA and Fama-French models, and they deliver low pricing errors for the FF5-alpha-weighted portfolio, whose average return cannot be fully explained by FF5 by construction. Importantly, the TS-alpha-PCA models outperform the PCA models because the time-series alpha target nudges the latent factors to capture risk premia information missed by FF5 and, in particular, to exploit long-short restrictions on characteristics not exploited by FF5.

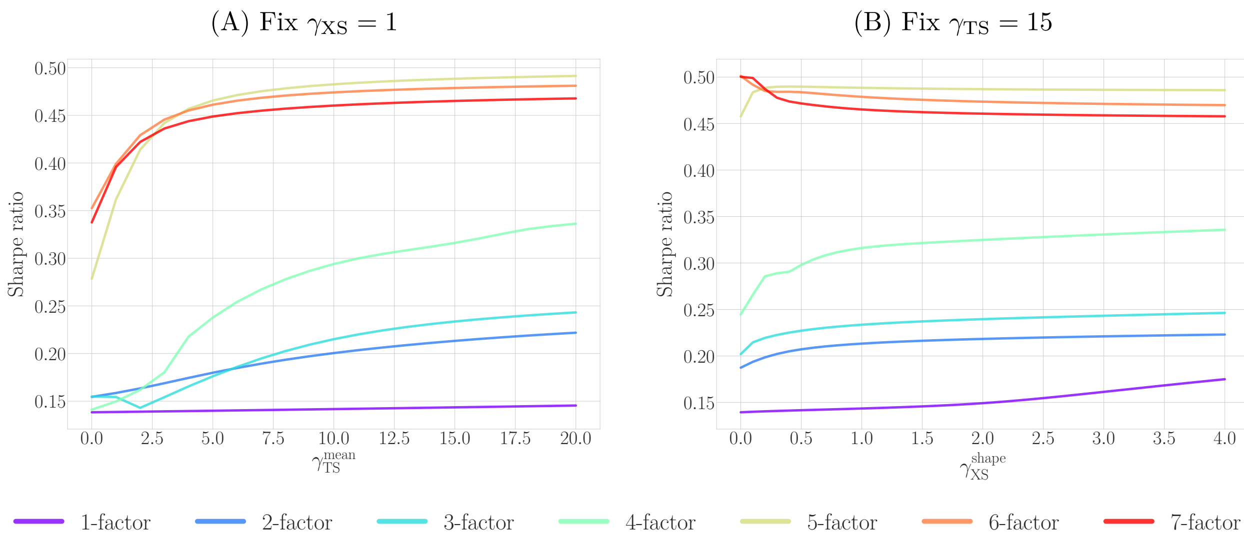
3.4 Combining cross-sectional and time-series targets

Having considered the models obtained by imposing the cross-sectional and time-series targets independently, we now investigate whether there is a benefit from considering both types of targets

[14] In Appendix C, we report the composition of the pure-alpha portfolio based on the PCA and TS-alpha-PCA models. The TS-alpha-PCA model identifies that the pure-alpha portfolio loads heavily on characteristics unexploited by FF5, such as those in the reversal and momentum categories.

Figure 9: Sharpe ratio for models that combine XS-shape and TS-mean targets

This figure illustrates how the out-of-sample Sharpe ratio of the cross-sectional time-series PCA models that combine both the shape and mean targets varies with the number of factors and the penalty parameters. Panel A illustrates the Sharpe ratio of each model as a function of the penalty on the mean target when the penalty on the shape target is $\gamma_{\text{XS}} = 1$ and Panel B as a function of the penalty on the shape target when the penalty on the mean target is $\gamma_{\text{TS}} = 15$.



simultaneously. In particular, we now consider a model that combines the cross-sectional shape and time-series mean targets.

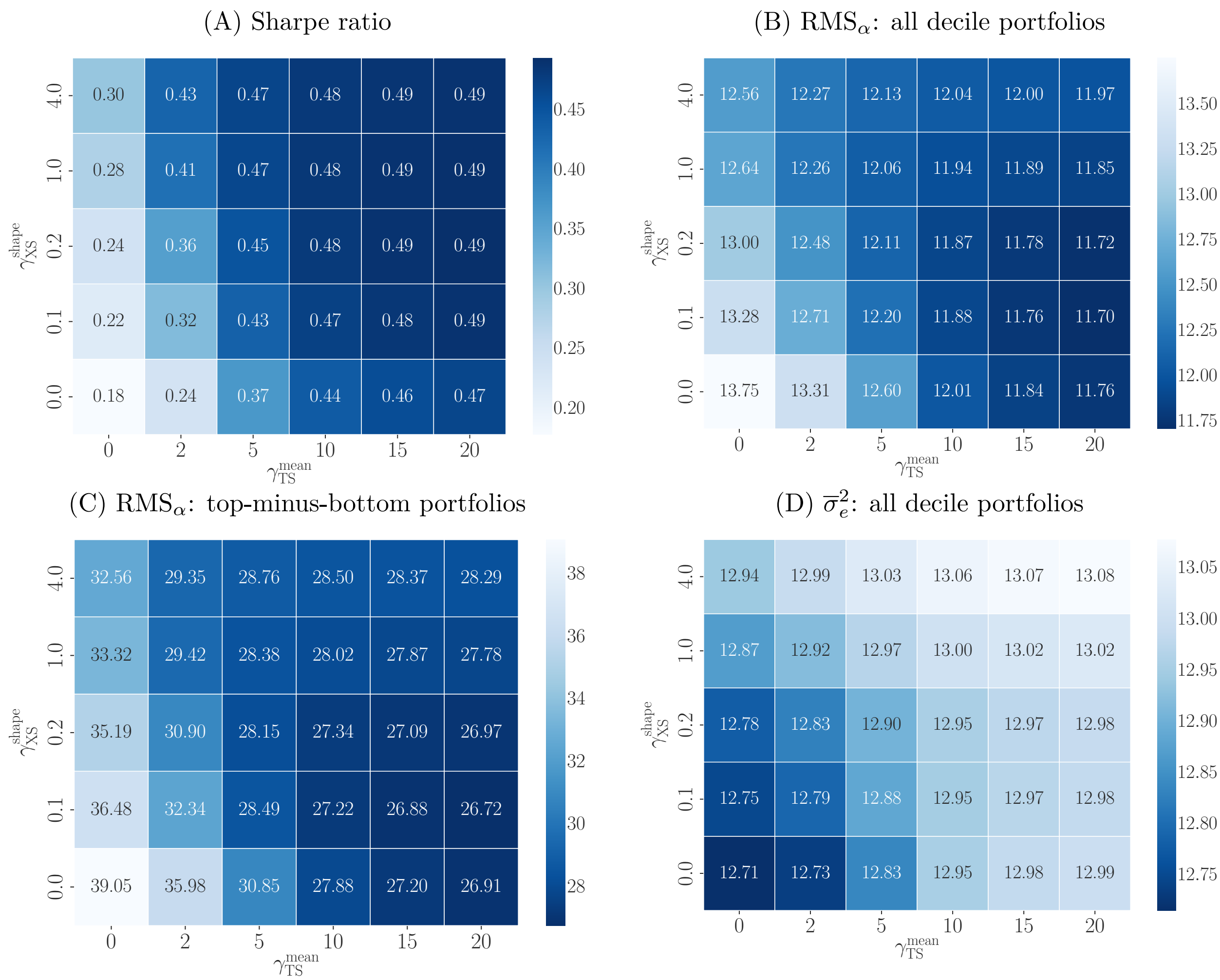
3.4.1 Number of factors and penalty parameters

Figure 9 illustrates how the out-of-sample Sharpe ratio of the cross-sectional time-series PCA models that combine both the shape and mean targets varies with the number of factors and the penalty parameters. Panel A illustrates the Sharpe ratio of each model as a function of the penalty on the mean target when the penalty on the shape target is $\gamma_{\text{XS}} = 1$, the benchmark penalty we use in Section 3.2.2, and Panel B as a function of the penalty on the shape target when the penalty on the mean target is $\gamma_{\text{TS}} = 15$, the benchmark penalty we use in Section 3.3.2.

Panel A of Figure 9 shows that, when we fix the penalty on the shape target to one, the Sharpe ratio increases substantially with the penalty on the mean target for every model with two or more factors. This is not surprising because the cross-sectional shape target is not explicitly designed to increase the Sharpe ratio of the model. To see this, note that even if the long-short characteristic portfolios capture risk premia information, the shape target does not help to identify *which* specific characteristics are most informative for the cross-section of expected returns. Thus, the time-series mean target complements the cross-sectional shape target by explicitly increasing the signal of characteristic long-short portfolios that carry high risk premia. Panel B shows that the cross-sectional shape target can also help to increase the Sharpe ratio of models with a time-

Figure 10: Performance of five-factor XS-shape-TS-mean-PCA models

This figure illustrates how the out-of-sample performance of the five-factor model varies with the penalty parameters on the cross-sectional shape and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, and average unexplained variance across the 370 decile portfolios, respectively. In each panel, the vertical axis depicts the penalty on the shape target and the horizontal axis the penalty on the mean target. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



series mean target, with the effect being more salient for low-dimensional models with at most five factors.

Figure 10 illustrates how the out-of-sample performance of the benchmark five-factor model varies with the penalty parameters on both the cross-sectional shape and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, and average unexplained variance across the 370 decile portfolios, respectively.[15] In each panel, the vertical and horizontal axes depict the penalty parameters on the shape and mean targets,

[15] Figure D.9 in Appendix D illustrates how the out-of-sample performance varies with the penalty parameters on both the cross-sectional shape and time-series mean targets for one- to six-factor models.

respectively. Figure 10 confirms the benefit of incorporating both types of targets. In particular, when the penalty on the mean target is low, the shape target helps to substantially improve the performance of the model. For example, when $\gamma_{\text{TS}} = 2$, as γ_{XS} increases from zero to one, the Sharpe ratio increases from 0.24 to 0.41, and the root-mean-squared pricing error for the top-minus-bottom portfolios decreases from 35.98 to 29.42.

In addition, Panel D of Figure 10 shows that simultaneously exploiting the two targets has a limited impact on the time-series explanatory power of the five-factor model. For example, in the extreme case in which $(\gamma_{\text{XS}}, \gamma_{\text{TS}}) = (4, 20)$, the average unexplained variance of the model is 13.08%, which is only slightly higher than that of the PCA model, 12.71%.

In summary, we find that one could benefit from exploiting both the shape and mean targets to improve the cross-sectional pricing performance without reducing the time-series explanatory power of the models. Consistent with the previous sections, in the next section we use $\gamma_{\text{XS}} = 1$ and $\gamma_{\text{TS}} = 15$ as the benchmark penalties for the XS-TS-PCA models.

3.4.2 Out-of-sample comparison

Table 4 compares the performance of the PCA and Fama-French models with that of five different cross-sectional times-series PCA models that exploit the following targets: (i) shape target (XS-shape-PCA), (ii) mean target (TS-mean-PCA), (iii) alpha target (TS-alpha-PCA), (iv) shape and mean targets combined (XS-shape-TS-mean-PCA), and (v) shape and alpha targets combined (XS-shape-TS-alpha-PCA). We consider models with three and five factors. Panels A and B contain the in-sample and out-of-sample results, respectively. In each panel, the first six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, pricing error for the FF5-alpha-weighted portfolio, and average unexplained variance across the 370 decile portfolios, respectively. The last two columns report the cross-sectional shape and time-series alpha metrics, respectively.

Table 4 shows that each target helps to improve performance in terms of the particular criterion they are designed to address. For instance, the PCA models explain most of the time-series variation of the returns. Models with a shape target deliver low pricing errors for the top-minus-bottom portfolios. Models with a mean target explain the 370 decile portfolios, and models with an alpha target explain the FF5-alpha-weighted portfolio.

Panel B in Table 4 shows that five-factor models with the targets substantially outperform three-factor models out of sample, which implies that the targets increase the signal of weak factors important for the cross-section and identify them as higher-order factors. For instance, the out-of-sample Sharpe ratio of the best-performing five-factor model (XS-shape-TS-mean-PCA) is 0.49, which is much larger than that of the best-performing three-factor model (XS-shape-TS-alpha-PCA), 0.28. Within five-factor models, we find that models that combine cross-sectional and time-series targets (XS-shape-TS-mean-PCA and XS-shape-TS-alpha-PCA) outperform other models, achieving higher Sharpe ratios and lower pricing errors across different test assets. Moreover,

Table 4: Performance of XS-shape-TS-mean-PCA and other models

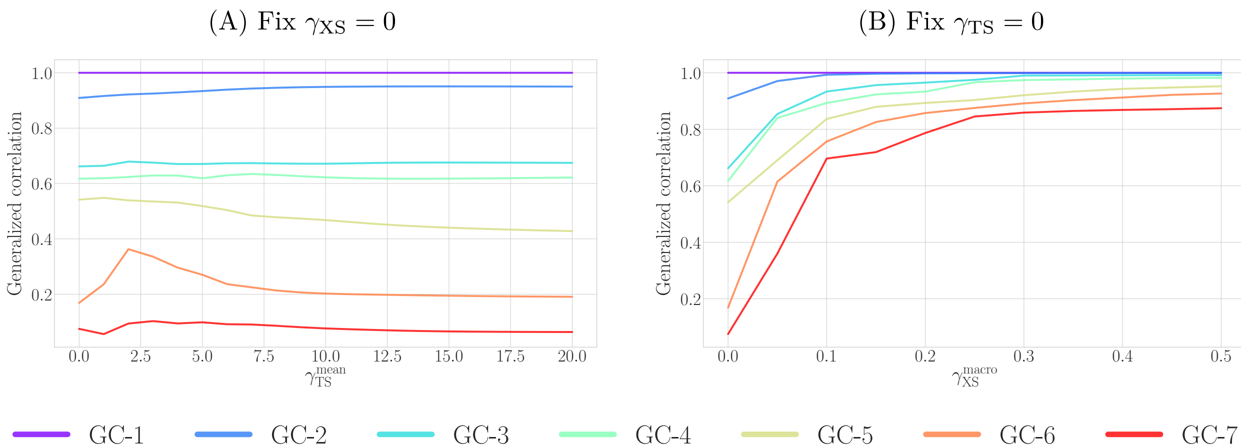
This table compares the performance of the PCA and Fama-French models with that of five different cross-sectional times-series PCA models that exploit the following targets: (i) shape target (XS-shape-PCA), (ii) mean target (TS-mean-PCA), (iii) alpha target (TS-alpha-PCA), (iv) shape and mean targets combined (XS-shape-TS-mean-PCA), and (v) shape and alpha targets combined (XS-shape-TS-alpha-PCA). We consider models with three and five factors. The penalty parameter on the cross-sectional targets is $\gamma_{\text{XS}} = 1$ and that on the time-series targets $\gamma_{\text{TS}} = 15$. Panels A and B contain the in-sample and out-of-sample results, respectively. In each panel, the first six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, pricing error for the FF5-alpha-weighted portfolio, and average unexplained variance across the 370 decile portfolios, respectively. The last two columns report the cross-sectional shape and time-series alpha metrics, respectively. Pricing errors are reported in basis points and unexplained variances in percentage. Bold numbers indicate the best-performing models.

Models	Sharpe ratio	RMS_α for 370 decile portfolios	RMS_α for TMB portfolios	Pricing error for FF5-alpha-weighted portfolio	$\overline{\sigma}_e^2$ for 370 decile portfolio	XS-shape metric	TS-alpha metric
			Panel A: in-sample				
Three-factor models							
FF	0.21	17.84	57.33	43.00	14.27	-	-
PCA	0.17	16.94	55.78	42.28	**12.57**	66.55	38.71
XS-shape-PCA	0.23	**14.77**	45.95	35.79	13.27	98.02	34.57
TS-mean-PCA	0.26	16.73	50.87	46.19	12.68	80.47	14.92
TS-alpha-PCA	0.16	18.28	58.88	32.98	12.62	71.65	**74.31**
XS-shape-TS-mean-PCA	**0.33**	15.50	**43.16**	32.93	13.31	98.42	53.75
XS-shape-TS-alpha-PCA	0.31	15.84	46.32	**24.04**	13.46	**98.97**	70.88
Five-factor models							
FF	0.32	16.09	49.45	46.22	13.27	-	-
PCA	0.25	14.23	42.89	34.01	**10.80**	80.99	49.14
XS-shape-PCA	0.33	13.38	38.35	32.52	10.91	97.87	46.80
TS-mean-PCA	0.57	13.44	31.88	15.66	10.95	86.02	88.37
TS-alpha-PCA	**0.66**	12.40	**26.48**	**4.42**	11.07	83.91	**99.35**
XS-shape-TS-mean-PCA	0.55	12.60	29.55	18.37	11.04	**98.63**	78.41
XS-shape-TS-alpha-PCA	0.62	**11.44**	27.26	8.07	11.14	98.40	89.12
			Panel B: out-of-sample				
Three-factor models							
FF	0.16	17.23	52.05	20.86	16.63	-	-
PCA	0.13	15.56	47.10	22.86	14.51	72.45	39.60
XS-shape-PCA	0.15	15.85	46.71	26.83	15.09	96.57	32.87
TS-mean-PCA	0.20	**14.13**	**38.99**	17.92	14.56	80.90	65.14
TS-alpha-PCA	0.27	15.15	44.16	**11.36**	**14.49**	76.81	**87.63**
XS-shape-TS-mean-PCA	0.23	14.85	40.01	22.52	15.01	**97.30**	50.66
XS-shape-TS-alpha-PCA	**0.28**	15.38	41.40	14.25	14.98	97.23	71.95
Five-factor models							
FF	0.31	13.66	36.92	28.41	15.64	-	-
PCA	0.18	13.75	39.05	20.45	**12.71**	77.97	45.29
XS-shape-PCA	0.28	12.64	33.32	20.58	12.87	96.69	41.08
TS-mean-PCA	0.46	**11.84**	**27.20**	10.23	12.98	76.91	86.22
TS-alpha-PCA	0.43	13.90	38.69	4.49	13.06	71.72	**99.37**
XS-shape-TS-mean-PCA	**0.49**	11.89	27.87	9.34	13.02	**97.35**	72.98
XS-shape-TS-alpha-PCA	0.48	13.56	37.48	**4.05**	13.12	96.56	84.96

when comparing these models with models that only exploit a time-series target (TS-mean-PCA and TS-alpha-PCA), we find that the models that exploit a cross-sectional target in addition to the time-series target have higher cross-sectional shape metrics. This suggests that their factor weights have a stronger long-short pattern on the characteristic decile portfolios. Therefore, the shape target not only transforms the space of returns spanned by the factors and improves the

Figure 11: Generalized correlations between macro mimicking portfolios and latent factors

Panels A and B graph the out-of-sample generalized correlations between the seven macro mimicking portfolios and the latent factors obtained with the TS-mean-PCA and XS-macro-PCA models, respectively. Both panels illustrate how the generalized correlations on the vertical axis vary with the penalty parameter on the horizontal axis.



cross-sectional pricing performance, but also facilitates the economic interpretation by nudging the factor weights to align with firm-specific characteristics.

3.5 Cross-sectional macro target

We now consider a *cross-sectional macro target* that exploits the economic insight of the intertemporal CAPM of Merton (1973) that asset exposure to innovations to macro variables that affect the marginal utility of consumption of a representative investor should be priced. Specifically, we consider the cross-sectional target matrix Λ^{XS} defined in (6), whose columns contain the covariances between the returns of each asset and the seven innovations to macroeconomic variables. Thus, the columns of Λ^{XS} are Fama-MacBeth-type mimicking portfolios for the innovations to the macroeconomic variables.

We construct the innovations to the macroeconomic variables based on the monthly data for 127 macroeconomic variables from FRED-MD following the procedure described in Section 3.1. In particular, we first transform these variables following the method suggested in the appendix of McCracken and Ng (2016) to render them stationary. Next, we standardize the transformed macroeconomic variables such that each of them has zero mean and unit standard deviation. We then apply PCA to these standardized variables and identify seven factors that drive their systematic variation. Lastly, we run a first-order vector autoregression on these PCA factors and take the residuals as the seven innovations to the macroeconomic variables.

3.5.1 Models with only macro target

Following Pelger (2020, II.C), we use generalized correlations[16] to study the time-series similarity between the mimicking portfolios for the seven macroeconomic innovations and the latent factors obtained with the TS-mean-PCA and XS-macro-PCA models. Recall that generalized correlations are the maximum correlation coefficients of linear combinations of two panels of factors.[17] Intuitively, the kth generalized correlation measures the time-series similarity between two factor models by considering the k-dimensional subspaces of the two models with the highest time-series "overlap." Therefore, if two sets of factors capture the same time-series variation, the generalized correlations between them are all equal to one. Panels A and B of Figure 11 graph the out-of-sample generalized correlations between the seven macro mimicking portfolios and the latent factors obtained with the TS-mean-PCA and XS-macro-PCA models, respectively.

Panel A in Figure 11 shows how the generalized correlations on the vertical axis vary with the penalty parameter on the time-series mean target on the horizontal axis. This panel shows that the PCA factors do not highly covary with the macro mimicking portfolios and the mean target does not help to increase covariation. For instance, the sixth and seventh generalized correlations between the macro mimicking portfolios and the latent factors are below 0.2 for PCA ($\gamma_{\text{TS}} = 0$) and also for TS-mean-PCA ($\gamma_{\text{TS}} = 15$). Therefore, neither the PCA factors nor the time-series mean PCA factors share the same time-series properties with the systematic risks related to macroeconomic fundamentals.

Panel B illustrates how the generalized correlations vary with the penalty parameter on the cross-sectional macro target. We find that the generalized correlations increase monotonically and substantially with the penalty on the macro target. For example, when γ_{XS} reaches 0.3, all seven generalized correlations are above 0.8. Therefore, the macro target nudges the latent factors to adopt the time-series properties of the macro mimicking portfolios.

Figure 12 illustrates how the out-of-sample performance of the cross-sectional macro PCA models varies with the number of factors and the penalty. In particular, Panel A gives a line plot depicting the Sharpe ratio of each model on the vertical axis as a function of the penalty on the horizontal axis. Panels B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, and average unexplained variance across the 370 decile portfolios, respectively. For each heatmap, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter.

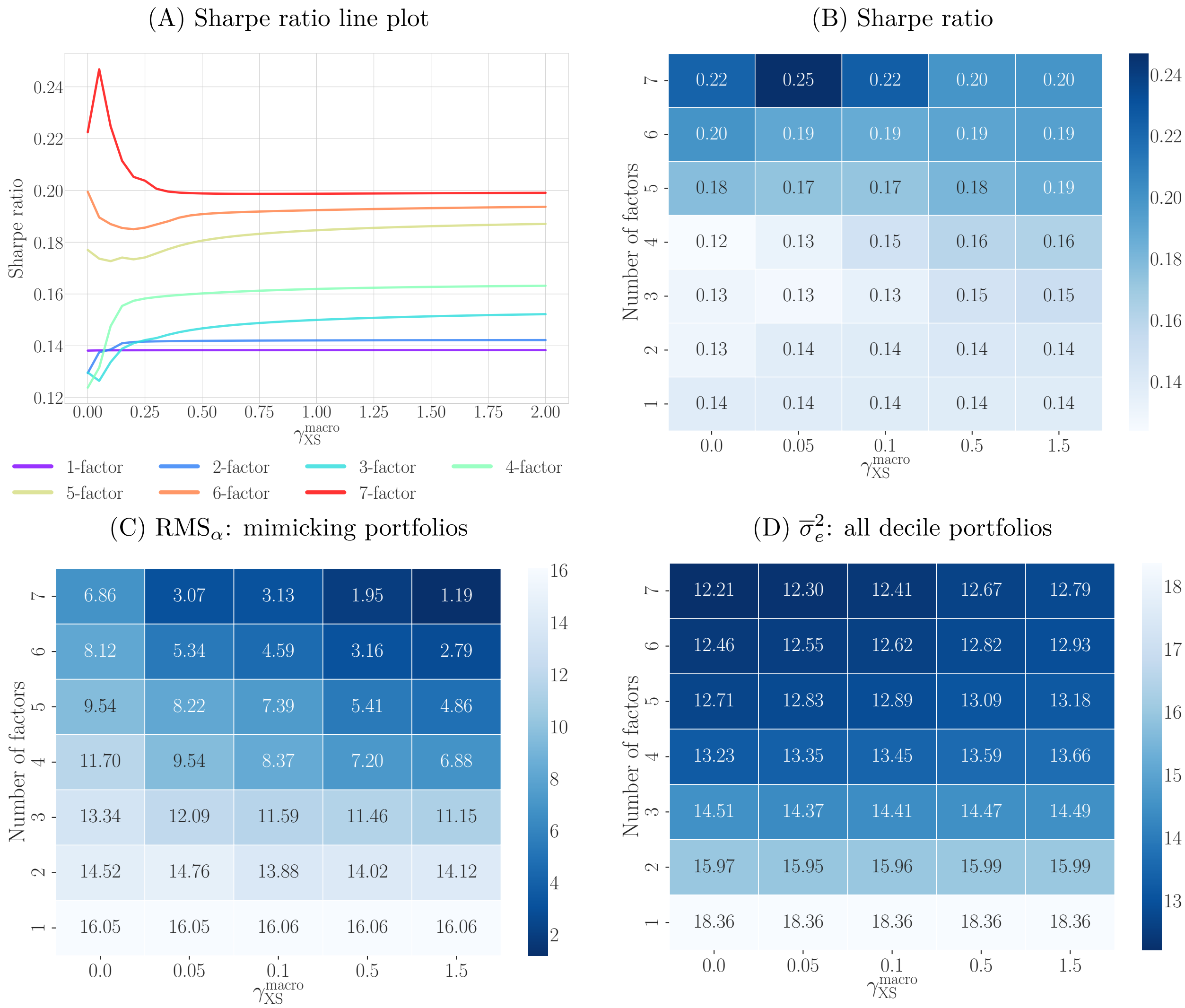
Panels A and B show that the macro target helps to increase the Sharpe ratio of low-dimensional models with two, three, and four factors. For example, the Sharpe ratio of the four-factor model increases from 0.12 to 0.16 when the penalty increases from zero (conventional PCA)

[16] Generalized correlation is also referred to as canonical correlation.

[17] Mathematically, to obtain the first generalized correlation between the m-dimensional random vector X and the n-dimensional random vector Y, one seeks vectors $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$ to maximize the correlation $\rho = \mathbb{C}\text{orr}(a^\top X, b^\top Y)$. By projecting $a^\top X$ out of X and $b^\top Y$ out of Y and repeating the above procedure, one obtains the second generalized correlation. This procedure can be repeated for $\min\{m, n\}$ times to obtain all generalized correlations $(\rho_1, \dots, \rho_{\min\{m,n\}})$ between X and Y.

Figure 12: Performance of cross-sectional macro PCA models

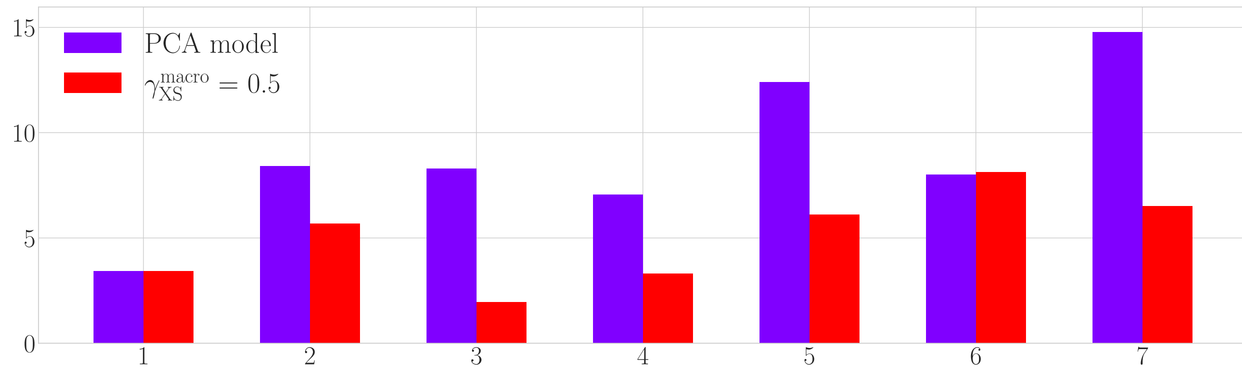
This figure illustrates the out-of-sample performance of cross-sectional macro PCA (XS-macro-PCA) models with up to seven factors and for different values of the penalty parameter. Panel A gives a line plot depicting the Sharpe ratio of each model on the vertical axis as a function of the penalty parameter, γ_{XS}, on the horizontal axis. Panels B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, and average unexplained variance across the 370 decile portfolios, respectively. For each heatmap, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, γ_{XS}. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



to 1.5 (XS-macro-PCA). However, the macro target has a limited impact on the Sharpe ratio of five- and six-factor models and decreases the Sharpe ratio of the seven-factor model. In particular, the Sharpe ratio of the six-factor model decreases from 0.20 to 0.19 when the penalty increases from zero to 0.5, a level for which the generalized correlation between the macro mimicking portfolios and the XS-macro-PCA factors is high. Thus, unlike the shape, mean, and alpha targets, the macro target does not help to identify assets with high risk premia.

Figure 13: Pricing errors for macro mimicking portfolios of PCA and XS-macro-PCA

This figure illustrates the out-of-sample pricing errors (in absolute value) for the seven macro mimicking portfolios of the five-factor PCA model (purple bars) and cross-sectional macro PCA model with $\gamma_{\mathrm{XS}} = 0.5$ (red bars). Pricing errors are reported in basis points.



Panel C in Figure 12 shows that the macro target helps to monotonically decrease the root-mean-squared pricing error across the seven macro mimicking portfolios. For example, the root-mean-squared pricing error for the five-factor model decreases by over 40% from 9.54 to 5.41 when the penalty increases from zero to 0.5. More specifically, Figure 13 compares the pricing errors (in absolute value) for the seven macro mimicking portfolios of the five-factor PCA model and the XS-macro-PCA model with $\gamma_{\mathrm{XS}} = 0.5$. We observe that the XS-macro-PCA model delivers *uniformly* lower pricing errors for the seven target mimicking portfolios. Thus, although the macro target does not help to identify assets with high risk premia, it does help to identify factors that price the macroeconomic mimicking portfolios better. Finally, this enhanced pricing ability does not come at the cost of lower time-series explanatory power. In particular, Panel D in Figure 12 shows that the average unexplained variance across all 370 decile portfolios does not vary substantially with the penalty on the macro target.

We now compare the performance of the XS-macro PCA model with that of the Fama-French and conventional PCA models. Panels A and B in Table 5 give the in-sample and out-of-sample results, respectively. In each panel, the first three columns report the acronym, Sharpe ratio, and root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, respectively. The last column reports each model's cross-sectional macro metric, which is calculated using Equation (20), where $P_{\Lambda}^{\mathrm{XS}}$ is the cross-sectional projection matrix associated with the macro target Λ^{XS} in Equation (6). The benchmark penalty is $\gamma_{\mathrm{XS}} = 0.5$, a level for which the XS-macro-PCA factors and the macro mimicking portfolios have high generalized correlations and the performance of the XS-macro-PCA model stabilizes.

Table 5 reveals a tradeoff between pricing the macro mimicking portfolios and pricing other assets with high risk premia. Specifically, the XS-macro-PCA models have high cross-sectional macro metrics (over 97 in all cases), and thus, their factors deliver low pricing errors for the macro mimicking portfolios. However, they completely ignore assets that are orthogonal to macroeco-

Table 5: Performance of cross-sectional macro PCA and other models

This table summarizes the performance of the Fama-French, PCA, and XS-macro-PCA models with three and five factors. The penalty parameter on the cross-sectional macro target is $\gamma_{\text{XS}} = 0.5$. Panels A and B report the in-sample and out-of-sample results, respectively. For each model, the four columns report the acronym, Sharpe ratio, root-mean-squared pricing error for the mimicking portfolios of the seven macroeconomic innovations, and cross-sectional macro metric, respectively. Root-mean-squared pricing errors are reported in basis points. Bold numbers indicate the best-performing models.

(A) In-sample

Models	Sharpe ratio	RMS_α for mimicking portfolios	XS-macro metric
Three-factor models			
FF	**0.21**	**25.55**	-
PCA	0.17	26.23	83.82
XS-PCA	0.17	26.91	**99.45**
Five-factor models			
FF	**0.32**	20.95	-
PCA	0.25	9.74	58.91
XS-PCA	0.24	**6.63**	**97.74**

(B) Out-of-sample

Models	Sharpe ratio	RMS_α for mimicking portfolios	XS-macro metric
Three-factor models			
FF	**0.16**	18.57	-
PCA	0.13	13.34	78.37
XS-PCA	0.15	**11.46**	**99.29**
Five-factor models			
FF	**0.31**	12.60	-
PCA	0.18	9.54	65.41
XS-PCA	0.18	**5.41**	**98.10**

nomic shocks but carry high risk premia, and thus, they achieve lower Sharpe ratios compared to the Fama-French models. Overall, although the macro target nudges the latent factors to reflect macroeconomic fundamental risks and explain the macroeconomic innovations, it does not help to price better a large cross-section of characteristic portfolios.

3.5.2 Combining macro and mean targets

We now consider models that combine the information of the cross-sectional macro and time-series mean targets.[18] Figure 14 illustrates how the out-of-sample Sharpe ratio of the models that combine the macro and mean targets varies with the number of factors and the penalty parameters. Panel A illustrates the Sharpe ratio of each model as a function of the penalty on the mean target when the penalty on the macro target is $\gamma_{\text{XS}} = 0.5$, the benchmark penalty we use in Section 3.5.1, and Panel B as a function of the penalty on the macro target when the penalty on the mean target is $\gamma_{\text{TS}} = 15$, the benchmark penalty we use in Section 3.3.2.

Panel A shows that, when we fix the penalty on the macro target to 0.5, the Sharpe ratio increases monotonically with the penalty on the mean target for models with two or more factors. This is expected because the time-series mean target explicitly increases the signal of factors with high risk premia. Panel B, on the other hand, shows that there is a tradeoff between pricing the macro mimicking portfolios and a broad cross section of characteristic decile portfolios. In particular, the Sharpe ratio of each model decreases substantially with the penalty on the macro target. For example, the Sharpe ratio of the five-factor model decreases by 55.2% from 0.46 to 0.21 when γ_{XS} increases from zero to two.

[18] We consider models with the cross-sectional macro and time-series alpha targets in the Internet Appendix.

Figure 14: Sharpe ratios for models with fixed γ_{XS} or γ_{TS}

This figure illustrates how the out-of-sample Sharpe ratio of cross-sectional time-series PCA models that combine both the macroeconomic and mean targets varies with the number of factors and the penalty parameters. Panel A illustrates the Sharpe ratio of each model as a function of the penalty on the time-series mean target when the penalty on the cross-sectional macro target is $\gamma_{\text{XS}} = 0.5$, and Panel B as a function of the penalty on the cross-sectional macro target when the penalty on the time-series mean target is $\gamma_{\text{TS}} = 15$.

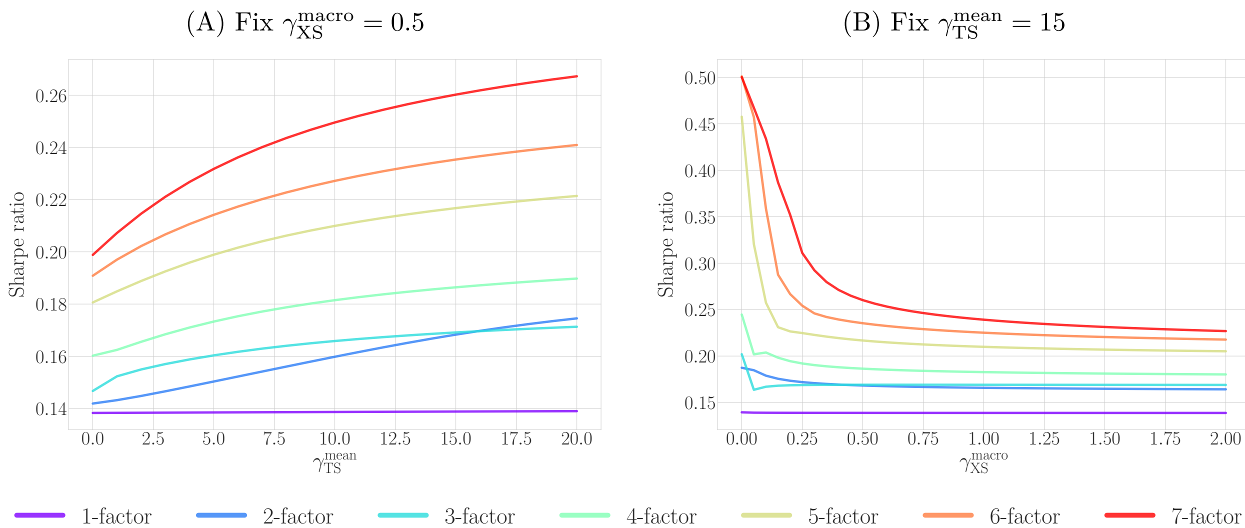


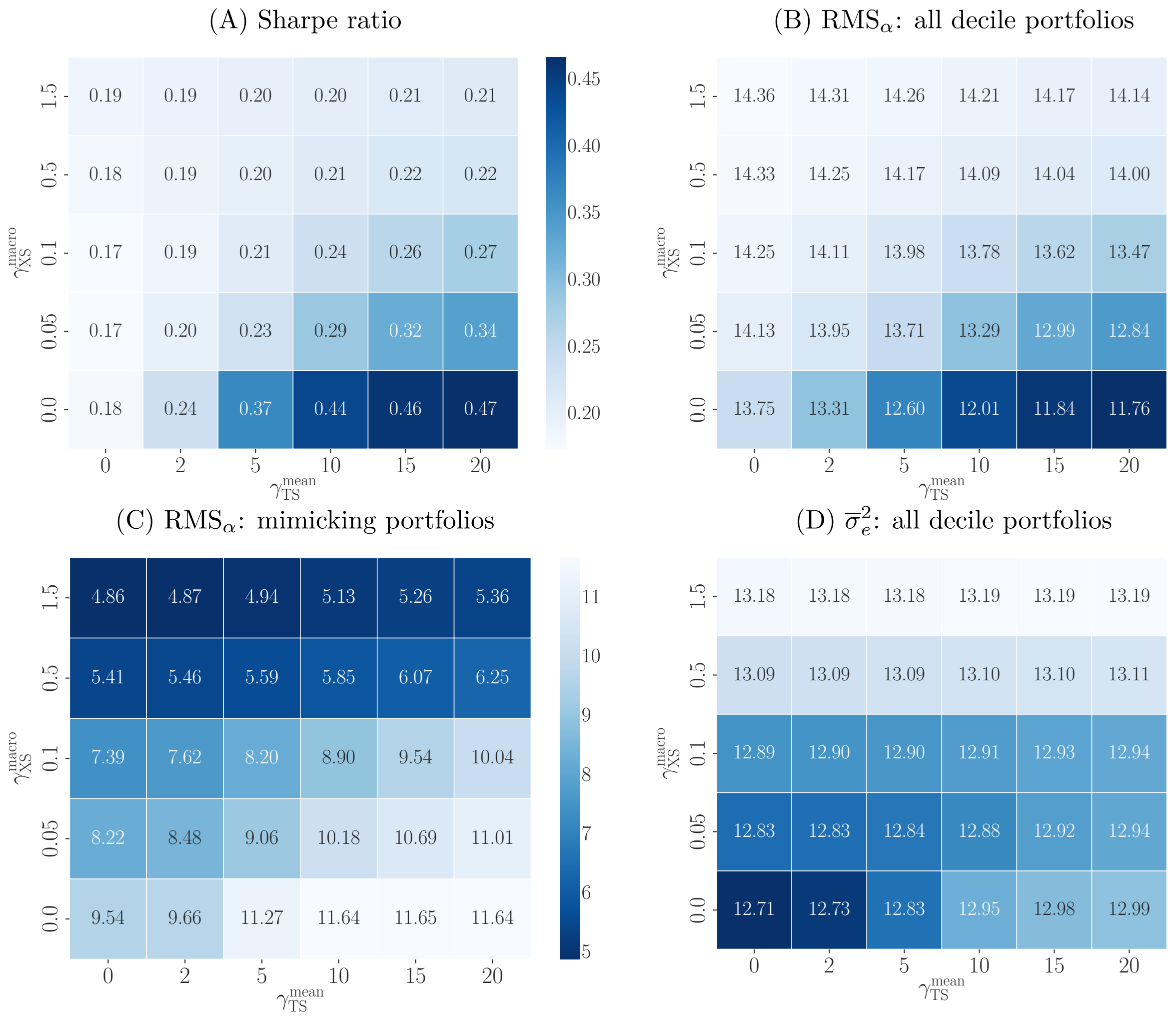
Figure 15 illustrates how the out-of-sample performance of the five-factor model varies with the penalty parameters on both the cross-sectional macro and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, and average unexplained variance across the 370 decile portfolios, respectively. In each panel, the vertical and horizontal axes depict the penalty parameters on the macro and mean targets, respectively.

Figure 15 confirms that there is a strong tradeoff between pricing a large cross-section of characteristic decile portfolios and pricing the macro mimicking portfolios. In particular, for a given penalty on the mean target, increasing the penalty on the macro target decreases the Sharpe ratio and increases the root-mean-squared pricing error across the 370 decile portfolios, but it decreases the pricing error for the seven macro mimicking portfolios. This suggests that the covariances between asset returns and macroeconomic innovations do not help to identify factors that explain the risk premia of the large cross-section of characteristic decile portfolios. In a nutshell, the information in macro variables does not help to explain a broad cross section of characteristic portfolios.

On the other hand, Figure 15 also shows that, for a given penalty on the macro target, increasing the penalty on the mean target increases the Sharpe ratio and decreases the root-mean-squared pricing error across all 370 decile portfolios, but it increases the root-mean-squared pricing error across the seven macro mimicking portfolios. In summary, while the mean target helps to

Figure 15: Performance of five-factor XS-macro-TS-mean-PCA models

This figure illustrates how the out-of-sample performance of the five-factor model varies with the penalty parameters on the cross-sectional macro and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, and average unexplained variance across the 370 decile portfolios, respectively. In each panel, the vertical vertical axis depicts the penalty on the macro target and the horizontal axis the penalty on the mean target. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



price a broad cross section of characteristic portfolios, it does not help to price the seven macro mimicking portfolios.

Table 6 compares the performance of the PCA model with that of three different cross-sectional time-series PCA models that exploit the following targets: (1) macro target (XS-macro-PCA), (ii) macro and mean targets combined (XS-macro-TS-mean-PCA), and (iii) macro and alpha targets combined (XS-macro-TS-alpha-PCA). We consider models with three and five factors. Panels A and B contain the in-sample and out-of-sample results, respectively. In each panel, the first six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the

Table 6: Performance of XS-macro-TS-mean-PCA and other models

This table summarizes the performance of three- and five-factor models including the PCA, XS-macro-PCA, XS-macro-TS-mean-PCA, and XS-macro-TS-alpha-PCA models. The benchmark penalty parameter on the cross-sectional target is $\gamma_{\text{XS}} = 0.5$ and that on the time-series target $\gamma_{\text{TS}} = 15$. Panels A and B contain the in-sample and out-of-sample results, respectively. For each model, the first six columns report the acronym, Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, pricing error for the FF5-alpha-weighted portfolio, and average unexplained variance across the 370 decile portfolios, respectively, and the last two columns report the cross-sectional macro and time-series alpha metrics, respectively. Pricing errors are reported in basis points and unexplained variances in percentage. Bold numbers indicate the best-performing models.

Models	Sharpe ratio	RMS_α for 370 decile portfolios	RMS_α for mimicking portfolios	Pricing error for FF5-alpha-weighted portfolio	$\overline{\sigma}_e^2$ for 370 decile portfolio	XS-macro metric	TS-alpha metric
Panel A: in-sample							
Three-factor models							
PCA	0.17	16.94	26.23	42.28	**12.57**	83.82	**38.71**
XS-macro-PCA	0.17	16.98	26.91	40.66	12.65	99.45	32.06
XS-macro-TS-mean-PCA	**0.18**	**16.73**	**25.97**	41.63	12.66	99.26	26.96
XS-macro-TS-alpha-PCA	0.16	17.08	27.82	**39.19**	12.66	**99.47**	37.70
Five-factor models							
PCA	0.25	**14.23**	9.74	34.01	**10.80**	58.91	**49.14**
XS-macro-PCA	0.24	14.38	**6.63**	33.91	11.01	**97.74**	39.12
XS-macro-TS-mean-PCA	**0.28**	14.24	9.31	32.19	11.02	96.77	43.46
XS-macro-TS-alpha-PCA	0.26	14.38	8.14	**31.33**	11.03	97.45	47.30
Panel B: out-of-sample							
Three-factor models							
PCA	0.13	15.56	13.34	22.86	14.51	78.37	39.60
XS-macro-PCA	0.15	15.39	11.46	22.88	14.47	**99.29**	35.62
XS-macro-TS-mean-PCA	**0.17**	**14.86**	10.53	22.32	14.47	98.73	39.29
XS-macro-TS-alpha-PCA	**0.17**	15.11	**9.45**	**20.13**	**14.39**	99.05	**44.92**
Five-factor models							
PCA	0.18	**13.75**	9.54	20.45	**12.71**	65.41	45.29
XS-macro-PCA	0.18	14.33	5.41	20.03	13.09	**98.10**	41.54
XS-macro-TS-mean-PCA	**0.22**	14.04	6.07	19.24	13.10	96.82	45.98
XS-macro-TS-alpha-PCA	0.21	14.40	**4.26**	**17.62**	13.10	97.57	**54.51**

370 decile portfolios, root-mean-squared pricing error across the seven macroeconomic mimicking portfolios, pricing error for the FF5-alpha-weighted portfolio, and average unexplained variance across the 370 decile portfolios, respectively. The last two columns report the cross-sectional macro and time-series alpha metrics, respectively.

Table 6 shows that each target helps to improve performance in terms of the particular criterion they are designed to address. For example, models with a time-series alpha target deliver the lowest pricing errors for the FF5-alpha-weighted portfolio. Different targets incorporate different economic information for the factor model to carry, and thus provide different objectives for the model to achieve. In some cases, the objectives of the cross-sectional and time-series targets can be achieved simultaneously, and thus assigning both targets with positive penalties helps to better achieve the common objective. For example, the XS-macro-TS-alpha-PCA models deliver the lowest out-of-sample root-mean-squared pricing errors for the seven macro mimicking portfolios, which implies that FF5 fails to fully capture the macroeconomic fundamental risk premia, and the macro mimicking portfolios have relatively high FF5 alphas. By assigning both the macro and alpha targets with positive penalties, the output model delivers low pricing errors for the macro mimicking portfolios.

However, in other cases, the objectives of the cross-sectional and time-series targets cannot be achieved simultaneously, and thus assigning both targets with positive penalties yields a tradeoff between the two different objectives. In particular, one could not find a model that simultaneously explains well the risk premia of the entire cross-section of characteristic decile portfolios and perfectly prices the low-premia macro mimicking portfolios. Therefore, the penalties on the macro and mean targets determine the relative weight of in achieving each of the objectives of the two targets.

4 Conclusion

This paper presents a general framework to encode researcher's beliefs about different properties of the risk factors, into their direct recovery from the cross-section of asset returns. Intuitively, imposing useful economic restrictions boosts the signal of weak factors and allows for a more efficient recovery of the underlying SDF. Our approach allows for a general type of both time-series and cross-sectional restrictions. In a large-dimensional empirical application we focus on investigating the role of four types of restrictions:

a) shape restrictions, related to the patterns of asset loading on the corresponding risk factors (empirically, we focus on the monotonicity of expected returns in decile-sorted portfolios);

b) spanning restriction, implying that recovered latent risk factors should (at least) reflect the pricing ability of the candidate exogenous state variables, e.g., macroeconomic innovations;

c) an APT-implied restriction that systematic risk factors should span the vector of expected returns (not just the sources of time-series variation);

d) explicit nudge of asset pricing factors towards spanning not just the main sources of time-series variation (which often coincides with a few popular reduced-form factors), but also the pricing errors generated by the latter, again with the intention to better span the underlying SDF and cross-section of asset returns.

We find that imposing informative economic restrictions and their combination can indeed be helpful in the recovery of systematic sources of risk, crucial for spanning a wide cross-section of asset returns. Empirically, this delivers risk factors that span a higher out-of-sample Sharpe ratio and yield lower pricing errors compared to the traditional approach. In doing so, we contribute to a large and growing literature that shows the importance of adapting statistical and machine-learning techniques by encoding suitable economic restrictions directly into the estimation procedure.

Our framework can be used to study a broad class of various asset pricing restrictions related to different spanning properties of the risk factors as well as shape restrictions on their loadings.

Importantly, we do not aim to provide a single most efficient way to recover the underlying SDF by choosing “optimal” priors. Instead, we allow the researcher to specify different types of restrictions consistent with both structural and reduced-form insights about the cross-section of asset returns and risk factors that drive it. In doing so, our framework can also be used in conjunction with other existing tools for estimating large-dimensional factor models, e.g., IPCA of Kelly et al. (2019). We leave the study of conditional factor models for individual stock returns, as well as further investigation of other economic restrictions, for future research.

References

Bai, Jushan, 2003, Inferential theory for factor models of large dimensions, *Econometrica* 71, 135–171.

Bai, Jushan, and Serena Ng, 2002, Determining the number of factors in approximate factor models, *Econometrica* 70, 191–221.

Bali, Tarun G., Amit Goyal, Dashan Huang, Fuwei Jiang, and Quan Wen, 2022, Predicting corporate bond returns: Merton meets machine learning, available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3686164`.

Bianchi, D., M. Buchner, and Andrea Tamoni, 2022, Bond risk premiums with machine learning, *Review of Financial Studies* 34, 1046–1089.

Blanchard, Olivier J., and Danny Quah, 1989, The dynamic effects of aggregate demand and supply disturbances, *American Economy Review* 79, 655–673.

Bryzgalova, Svetlana, Jiantao Huang, and Christian Julliard, 2022, Bayesian solutions for the factor zoo: We just ran two quadrillion models, *Journal of Finance* Forthcoming.

Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica: Journal of the Econometric Society* 1281–1304.

Chen, Luyang, Markus Pelger, and Jason Zhu, 2020, Deep learning in asset pricing, *Available at SSRN* 3350138.

Chen, Nai-Fu, Richard Roll, and Stephen A. Ross, 1986, Economic forces and the stock market, *Journal of Business* 59, 383–403.

Connor, Gregory, and Robert A. Korajczyk, 1986, Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics* 15, 373–394.

Connor, Gregory, and Robert A Korajczyk, 1988, Risk and return in an equilibrium apt: Application of a new test methodology, *Journal of financial economics* 21, 255–289.

Dello Preite, Masimo, Raman Uppal, Paolo Zaffaroni, and Irina Zviadadze, 2022, What is missing in asset-pricing factor models?, available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4135146`.

Dew-Becker, Ian, and Stefano Giglio, 2016, Asset Pricing in the Frequency Domain: Theory and Empirics, *The Review of Financial Studies* 29, 2029–2068.

Duan, Junting, Markus Pelger, and Ruoxuan Xiong, 2022, Target-pca: Transfer learning large dimensional panel data, *Working paper* .

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fan, Jianqing, Yuan Liao, and Weichen Wang, 2016, Projected principal component analysis in factor models, *Annals of Statistics* 44, 219–254.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020a, Taming the factor zoo: A test of new factors, *Journal of Finance* 75, 1327–1370.

Feng, Guanhao, Nick Polson, and Jianeng Xu, 2020b, Deep learning in characteristics-sorted factor models, *Available at SSRN* 3243683.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2016, Dissecting characateristics non-parametrically, Chicago Booth School of Business working paper.

Freyberger, Joachim, Andreas Neuhierl, and Michael Weber, 2020, Dissecting characteristics non-parametrically, *Review of Financial Studies* 33, 2326–2377.

Giglio, Stefano, and Dacheng Xiu, 2021, Asset pricing with omitted factors, *Journal of Political Economy* 129.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2018, Empirical asset pricing via machine learning, Technical report, National Bureau of Economic Research.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2021, Autoencoder asset pricing models, *Journal of Econometrics* 222, 429–450.

He, Zhiguo, Bryan Kelly, and Asaf Manela, 2017, Intermediary asset pricing: New evidence from many asset classes, *Journal of Financial Economics* 126, 1–35.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: An investment approach, *Review of Financial Studies* 28, 650–705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2018, Replicating Anomalies, *The Review of Financial Studies* 33, 2019–2133.

Huberman, Gur, 1982, A simple approach to the arbitrage pricing theory, *Journal of Economic Theory* 28, 183–191.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jensen, Theis I., Bryan T. Kelly, and Lasse H. Pedersen, 2022, Is there a replication crisis in finance?, *Journal of finance* Forthcoming.

Kelly, Bryan T, Seth Pruitt, and Yinan Su, 2019, Characteristics are covariances: A unified model of risk and return, *Journal of Financial Economics* 134, 501–524.

Kim, Soohun, Robert A Korajczyk, and Andreas Neuhierl, 2020, Arbitrage Portfolios, *The Review of Financial Studies* 34, 2813–2856.

Kozak, Serhiy, Stefan Nagel, and Shrihari Santosh, 2020, Shrinking the cross-section, *Journal of Financial Economics* 135, 271–292.

Lettau, Martin, and Sydney Ludvigson, 2001, Resurrecting the (C)CAPM: A cross-sectional test when risk premia are time-varying, *Journal of Political Economy* 109, 1238–1287.

Lettau, Martin, and Markus Pelger, 2020a, Estimating latent asset-pricing factors, *Journal of Econometrics* 218, 1–31.

Lettau, Martin, and Markus Pelger, 2020b, Factors that fit the time series and cross-section of stock returns, *Review of Financial Studies* 33, 2274–2325.

Lintner, John, 1965, Security, risk, and maximal gains from diversification, *Journal of Finance* 20, 587–615.

McCracken, Michael W., and Serena Ng, 2016, FRED-MD: A monthly database for macroeconomic research, *Journal of Business & Economic Statistics* 34, 574–589.

Merton, Robert C, 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Neuhierl, Andreas, and Rasmus T. Varneskov, 2021, Frequency dependent risk, *Journal of Financial Economics* 140, 644–675.

Onatski, Alexey, 2012, Asymptotics of the principal components estimator of large factor models with weakly influential factors, *Journal of Econometrics* 168, 244–258.

Patton, Andrew J., and Allan Timmermann, 2010, Monotonicity in asset returns: New tests with applications to the term structure, the capm, and portfolio sorts, *Journal of Financial Economics* 98, 605–625.

Pelger, Markus, 2020, Understanding systematic risk: a high-frequency approach, *The Journal of Finance* 75, 2179–2220.

Raponi, Valentina, Cesare Robotti, and Paolo Zaffaroni, 2019, Testing Beta-Pricing Models Using Large Cross-Sections, *The Review of Financial Studies* 33, 2796–2842.

Ross, Stephen, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium condition of risk, *Journal of Finance* 19, 425–442.

Stock, James H, and Mark W Watson, 2002, Forecasting using principal components from a large number of predictors, *Journal of the American statistical association* 97, 1167–1179.

Uhlig, Harald, 2005, What are the effects of monetary policy on output? results from an agnostic identification procedure, *Journal of Monetary Economics* 52, 381–419.

Uppal, Raman, Paolo Zaffaroni, and Irina Zviadadze, 2021, Correcting misspecified stochastic discount factors, working paper.

A Methodology

A.1 General Statement

Proposition A.1 *The XS-TS-Target-PCA factors can be obtained from PCA applied in the time- or cross-sectional dimension.*

1. *The latent-factor model that optimizes the XS-TS-Target-PCA objective in* (11) *can be equivalently obtained by:*

   (a) *Applying PCA to the matrix*

$$\frac{1}{NT}\left(I_N + \tilde{\gamma}_{XS} P_{\Lambda}^{XS}\right) R^{\top} \left(I_T + \gamma_{TS} P_G^{TS}\right) R \left(I_N + \tilde{\gamma}_{XS} P_{\Lambda}^{XS}\right), \tag{A1}$$

   where $\tilde{\gamma}_{XS} = \sqrt{\gamma_{XS} N/L + 1} - 1$*. Specifically, let* $Q_K \in \mathbb{R}^{N \times K}$ *be matrix whose columns contain the first* K *eigenvectors of matrix* (A1)*, then the estimated factor loadings and factor returns are*

$$\hat{\beta}_{XS\text{-}TS} = \left[\left(\tilde{\gamma}_{XS} + 1\right) I_N - \tilde{\gamma}_{XS} P_{\Lambda}^{XS}\right] Q_K, \qquad \textit{and} \tag{A2}$$

$$\hat{F}_{XS\text{-}TS} = R \left(\frac{1}{\tilde{\gamma}_{XS} + 1} I_N + \frac{\tilde{\gamma}_{XS}}{\tilde{\gamma}_{XS} + 1} P_{\Lambda}^{XS}\right) Q_K, \quad \textit{respectively.} \tag{A3}$$

   (b) *Applying PCA to the matrix*

$$\frac{1}{NT}\left(I_T + \tilde{\gamma}_{TS} P_G^{TS}\right) R \left(I_N + \frac{\gamma_{XS} N}{L} P_{\Lambda}\right) R^{\top} \left(I_N + \tilde{\gamma}_{TS} P_G^{TS}\right), \tag{A4}$$

   where $\tilde{\gamma}_{TS} = \sqrt{\gamma_{TS} + 1} - 1$*. Specifically, let* $U_K \in \mathbb{R}^{T \times K}$ *be matrix whose columns contain the first* K *eigenvectors of matrix* (A4)*, then the estimated factor returns and factor loadings are*

$$\hat{F}_{XS\text{-}TS} = \left[\left(\tilde{\gamma}_{TS} + 1\right) I_T - \tilde{\gamma}_{TS} P_G^{TS}\right] U_K, \qquad \textit{and} \tag{A5}$$

$$\hat{\beta}_{XS\text{-}TS} = R^{\top} \left(\frac{1}{\tilde{\gamma}_{TS} + 1} I_T + \frac{\tilde{\gamma}_{TS}}{\tilde{\gamma}_{TS} + 1} P_G^{TS}\right) U_K, \quad \textit{respectively.} \tag{A6}$$

2. *In the special case of XS-Target-PCA, the latent-factor model that optimizes the XS-Target-PCA objective in* (4) *can be equivalently obtained by:*

   (a) *Applying PCA to the matrix*

$$\frac{1}{NT}\left(I_N + \tilde{\gamma}_{XS} P_{\Lambda}^{XS}\right) R^{\top} R \left(I_N + \tilde{\gamma}_{XS} P_{\Lambda}^{XS}\right) \in \mathbb{R}^{N \times N}, \tag{A7}$$

where $\tilde{\gamma}_{XS} = \sqrt{\gamma_{XS} N/L + 1} - 1$. *Specifically, let* $Q_K \in \mathbb{R}^{N \times K}$ *be the matrix whose columns contain the first* K *eigenvectors of matrix* (A7)*, then the estimated factor loadings and factor returns are*

$$\hat{\beta}_{XS} = \left[(\tilde{\gamma}_{XS} + 1) I_N - \tilde{\gamma}_{XS} P_{\Lambda}^{XS} \right] Q_K, \qquad \textit{and} \tag{A8}$$

$$\hat{F}_{XS} = R \left(\frac{1}{\tilde{\gamma}_{XS} + 1} I_N + \frac{\tilde{\gamma}_{XS}}{\tilde{\gamma}_{XS} + 1} P_{\Lambda}^{XS} \right) Q_K, \quad \textit{respectively.} \tag{A9}$$

(b) *Applying PCA to the matrix*

$$\frac{1}{NT} R \left(I_N + \frac{\gamma_{XS} N}{L} P_{\Lambda}^{XS} \right) R^{\top} \in \mathbb{R}^{T \times T}. \tag{A10}$$

Specifically, the estimated factor returns, $\hat{F}_{XS}$*, are the first* K *eigenvectors of matrix* (A10)*, and the estimated factor loadings are* $\hat{\beta}_{XS} = R^{\top} \hat{F}_{XS}$*.*

3. *In the special case of TS-Target-PCA, The latent-factor model that optimizes the TS-PCA objective in* (8) *can be equivalently obtained by:*

   (a) *Applying PCA to the matrix*

$$\frac{1}{NT} R^{\top} \left(I_T + \gamma_{TS} P_G^{TS} \right) R \in \mathbb{R}^{N \times N}. \tag{A11}$$

Specifically, the estimated factor loadings $\hat{\beta}_{TS}$ *are the first* K *eigenvectors of matrix* (A11)*, and the factor returns are* $\hat{F}_{TS} = R \hat{\beta}_{TS}$*.*

   (b) *Applying PCA to the matrix*

$$\frac{1}{NT} \left(I_T + \tilde{\gamma}_{TS} P_G^{TS} \right) R R^{\top} \left(I_T + \tilde{\gamma}_{TS} P_G^{TS} \right) \in \mathbb{R}^{T \times T}, \tag{A12}$$

where $\tilde{\gamma}_{TS} = \sqrt{\gamma_{TS} + 1} - 1$. *Specifically, let* $U_K \in \mathbb{R}^{T \times K}$ *be the matrix whose columns contain the first* K *eigenvectors of matrix* (A12)*, then the estimated factor returns and factor loadings are*

$$\hat{F}_{TS} = \left[(\tilde{\gamma}_{TS} + 1) I_T - \tilde{\gamma}_{TS} P_G^{TS} \right] U_K, \qquad \textit{and} \tag{A13}$$

$$\hat{\beta}_{TS} = R^{\top} \left(\frac{1}{\tilde{\gamma}_{TS} + 1} I_T + \frac{\tilde{\gamma}_{TS}}{\tilde{\gamma}_{TS} + 1} P_G^{TS} \right) U_K, \quad \textit{respectively.} \tag{A14}$$

A.2 Proof of all results

This section contains the proof of the second part of Proposition A.1. Proofs for the other two parts are similar and thus are omitted to avoid repetition.

Part 1:

The first-order condition of problem (4) with respect to F is

$$-\frac{2}{NT}\left(R-F\beta^{\top}\right)\beta-\frac{2\gamma_{\mathrm{XS}}}{LT}\left(RP_{\Lambda}^{\mathrm{XS}}-F\beta^{\top}P_{\Lambda}^{\mathrm{XS}}\right)P_{\Lambda}^{\mathrm{XS}}\beta=0$$

$$\Rightarrow\quad F=R\left(I_N+\frac{\gamma_{\mathrm{XS}}N}{L}P_{\Lambda}^{\mathrm{XS}}\right)\beta\left[\beta^{\top}\left(I_N+\frac{\gamma_{\mathrm{XS}}N}{L}P_{\Lambda}^{\mathrm{XS}}\right)\beta\right]^{-1}. \tag{A15}$$

Let

$$\tilde{\beta}=\frac{1}{\tilde{\gamma}_{\mathrm{XS}}+1}\left(I_N+\tilde{\gamma}_{\mathrm{XS}}P_{\Lambda}\right)\beta, \tag{A16}$$

where $\tilde{\gamma}_{\mathrm{XS}}=\sqrt{\gamma_{\mathrm{XS}}N/L+1}-1$. Note that problem (4) is equivalent to

$$\begin{aligned}
&\min_{F,\beta}\quad \operatorname{trace}\left[\left(I_N+\frac{\gamma_{\mathrm{XS}}N}{L}P_{\Lambda}^{\mathrm{XS}}\right)\left(R^{\top}R-R^{\top}F\beta^{\top}-\beta F^{\top}R+\beta F^{\top}F\beta\right)\right]\\
\Leftrightarrow\quad &\max_{F,\beta}\quad \operatorname{trace}\left[\left(I_N+\frac{\gamma_{\mathrm{XS}}N}{L}P_{\Lambda}^{\mathrm{XS}}\right)\left(R^{\top}F\beta^{\top}+\beta F^{\top}R-\beta F^{\top}F\beta\right)\right].
\end{aligned} \tag{A17}$$

Plugging equations (A15) and (A16) into (A17), we obtain that problem (4) becomes

$$\max_{\tilde{\beta}}\quad \operatorname{trace}\left[\left(\tilde{\beta}^{\top}\tilde{\beta}\right)^{-1}\tilde{\beta}^{\top}\left(I_N+\tilde{\gamma}_{\mathrm{XS}}P_{\Lambda}\right)R^{\top}R\left(I_N+\tilde{\gamma}_{\mathrm{XS}}P_{\Lambda}\right)\tilde{\beta}\right]. \tag{A18}$$

Let $Q_K\in\mathbb{R}^{N\times K}$ be the matrix whose columns contain the first K eigenvectors of matrix (A7). If we normalize $\tilde{\beta}^{\top}\tilde{\beta}=I_K$, then $\tilde{\beta}$ that solves problem (A18) is Q_K. Plugging $\tilde{\beta}=Q_K$ into equations (A16) and (A15), we obtain

$$\hat{\beta}_{\mathrm{XS}}=\left[\left(\tilde{\gamma}_{\mathrm{XS}}+1\right)I_N-\tilde{\gamma}_{\mathrm{XS}}P_{\Lambda}^{\mathrm{XS}}\right]Q_K,\qquad\text{and} \tag{A19}$$

$$\hat{F}_{\mathrm{XS}}=R\left(\frac{1}{\tilde{\gamma}_{\mathrm{XS}}+1}I_N+\frac{\tilde{\gamma}_{\mathrm{XS}}}{\tilde{\gamma}_{\mathrm{XS}}+1}P_{\Lambda}^{\mathrm{XS}}\right)Q_K,\quad\text{respectively.} \tag{A20}$$

Part 2:

The first-order condition of problem (4) with respect to β is

$$-\frac{2}{NT}\left(R^{\top}-\beta F^{\top}\right)F-\frac{2\gamma_{\mathrm{XS}}}{LT}P_{\Lambda}^{\mathrm{XS}}\left(P_{\Lambda}^{\mathrm{XS}}R^{\top}-P_{\Lambda}^{\mathrm{XS}}\beta F^{\top}\right)F=0$$

$$\Rightarrow \quad \beta = R^{\top} F \left(F^{\top} F \right)^{-1}. \tag{A21}$$

Plugging equation (A21) into (A17), we obtain that problem (4) becomes

$$\max_{F} \quad \text{trace} \left[\left(F^{\top} F \right)^{-1} F^{\top} R \left(I_N + \frac{\gamma_{\text{XS}} N}{L} P_{\Lambda}^{\text{XS}} \right) R^{\top} F \right]. \tag{A22}$$

Let $U_K \in \mathbb{R}^{T \times K}$ be the matrix that contain the first K eigenvectors of matrix (A10). If we normalize $F^{\top} F = I_K$, then F that solves problem (A22) is $\hat{F}_{\text{XS}} = U_K$. Plugging $\hat{F}_{\text{XS}}$ into equation (A21), we obtain $\hat{\beta}_{\text{XS}} = R^{\top} \hat{F}_{\text{XS}}$.

B Firm-specific characteristics

Table B.1 classifies the 37 firm-specific characteristics in our dataset into eight categories. The first column reports the name of each category, the second column the acronyms of characteristics in each category, and the third (fourth) column the average return (in percentage) of the bottom-decile (top-decile) portfolios in each category. The last column reports the average Sharpe ratio of the top-minus-bottom portfolios, which are constructed by going long the top-decile and short the bottom-decile portfolios, of each category. Table B.1 shows that top-minus-bottom portfolios in the reversal and value interaction categories earn the highest Sharpe ratios.

Table B.1: Categories of characteristics

This table classifies the 37 characteristics into eight categories. The first column reports the name of each category, the second column reports the acronyms of the characteristics in each category, the third (fourth) column reports the average return (in percentage) of the bottom-decile (top-decile) portfolios in each category, and the last column reports the average Sharpe ratio of the top-minus-bottom portfolios for characteristics in each category.

Category	Characteristics	Mean bottom-decile	Mean top-decile	SR top-minus-bottom
reversal	*lrrev, strev, indmomrev, indrrev, indrrevlv*	0.18	1.02	0.23
value interaction	*valmom, valmomprof, valprof*	0.34	1.03	0.16
investment	*inv, invcap, igrowth, growth, noa*	0.39	0.72	0.10
momentum	*mom, mom12, indmom, momrev*	0.30	0.93	0.10
value	*value, valuem, divp, ep, cfp, sp*	0.44	0.87	0.09
other	*size, price, accruals, ciss, gmargins, lev season, sgrowth*	0.42	0.69	0.07
trading frictions	*ivol, shvol, aturnover*	0.26	0.57	0.06
profitability	*prof, roaa, roea*	0.42	0.64	0.06

C Weights of pure-alpha portfolio

Panels A and B in Figure C.1 illustrate the weights of the pure-alpha portfolio on the extreme-decile portfolios based on the five-factor TS-alpha-PCA model with $\gamma_{\text{TS}}^{\text{alpha}} = 15$ and the five-factor PCA model, respectively. In particular, we use the regression coefficients in Table 3 to obtain these weights. Specifically, for the TS-alpha-PCA model, the weights are

$$0.00 \times \text{TS-alpha}_1 - 0.33 \times \text{TS-alpha}_2 + 0.24 \times \text{TS-alpha}_3 + 0.47 \times \text{TS-alpha}_4 + 0.33 \times \text{TS-alpha}_5,$$

where TS-alpha$_i$ is the weights on the extreme-decile portfolios of the ith TS-alpha-PCA factor. For the PCA model, the weights are

$$-0.00 \times \text{PCA}_1 - 0.05 \times \text{PCA}_2 - 0.09 \times \text{PCA}_3 + 0.28 \times \text{PCA}_4 + 0.10 \times \text{PCA}_5,$$

where PCA$_i$ is the weights on the extreme-decile portfolios of the ith PCA factor.

Panel A shows that based on the TS-alpha-PCA model, the pure-alpha portfolio assigns high weights on long-short portfolios of characteristics in the reversal, value interaction, momentum, and other categories, none of which is exploited by FF5. For characteristics exploited by FF5, the pure-alpha portfolio has negligible weights on extreme-decile portfolios of the value and investment characteristics, and only has positive weights on the high-return top-decile portfolios in the profitability category.

Note that the TS-alpha-PCA factors well explain the expected return of the pure-alpha portfolio. Therefore, the TS-alpha-PCA factors identify that the risk premia FF5 fails to capture are carried by the long-short portfolios of characteristics in the reversal, value interaction, momentum, and other categories.

Panel B shows that based on the PCA model, the pure-alpha portfolio does not load heavily on the extreme-decile portfolios. In particular, it does not load on the reversal portfolios, which have the highest Sharpe ratios. Therefore, the reason why PCA factors fail to explain the expected return of pure-alpha portfolio is that they do not capture risk premia of long-short portfolios of characteristics unexploited by FF5, which the pure-alpha portfolio carry.

Figure C.1: Weights of pure-alpha portfolio on extreme-decile portfolios based on PCA and TS-alpha-PCA factors

Panels A and B illustrate the weights of the pure-alpha portfolio on the extreme-decile portfolios based on the five-factor TS-alpha-PCA model with $\gamma_{\mathrm{TS}}^{\mathrm{alpha}} = 15$ and the five-factor PCA model, respectively. In each panel, the horizontal axis depicts the characteristic categories. Table 3 reports the regression coefficients to obtain the weights of the pure-alpha portfolio. Specifically, for the TS-alpha-PCA model, the weights are

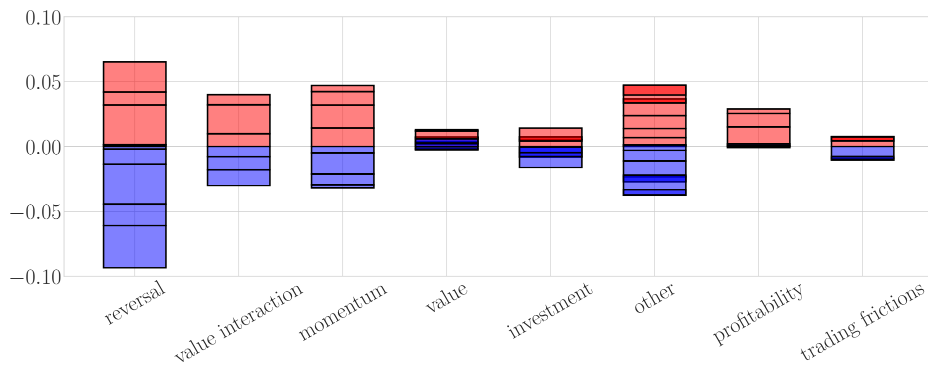
$$0.00 \times \text{TS-alpha}_1 - 0.33 \times \text{TS-alpha}_2 + 0.24 \times \text{TS-alpha}_3 + 0.47 \times \text{TS-alpha}_4 + 0.33 \times \text{TS-alpha}_5,$$

where TS-alpha$_i$ is the weights on the extreme-decile portfolios of the ith TS-alpha-PCA factor. For the PCA model, the weights are

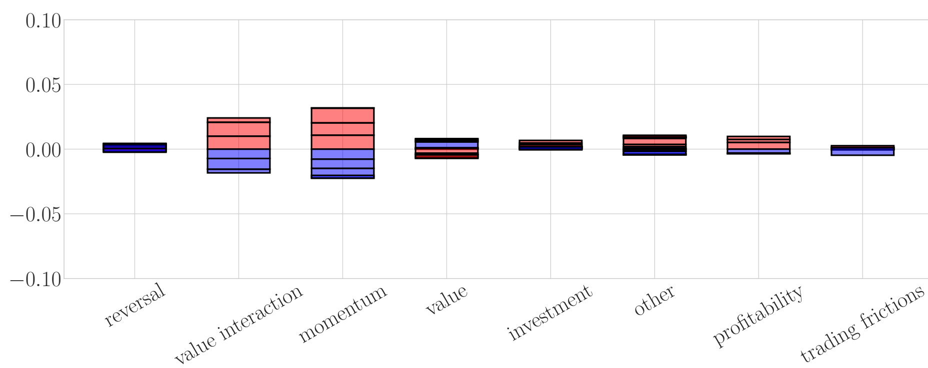
$$-0.00 \times \mathrm{PCA}_1 - 0.05 \times \mathrm{PCA}_2 - 0.09 \times \mathrm{PCA}_3 + 0.28 \times \mathrm{PCA}_4 + 0.10 \times \mathrm{PCA}_5,$$

where PCA$_i$ is the weights on the extreme-decile portfolios of ith PCA factor. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) TS-alpha-PCA



(B) PCA



D Additional figures

This section contains additional figures to illustrate the in-sample factor weights and out-of-sample performance of different models. Section D.1 includes figures that illustrate the in-sample factor weights on the decile portfolios of models that exploit different cross-sectional and time-series targets. Section D.2 includes figures that report the out-of-sample performance of models that exploit the following targets: (i) time-series mean target (TS-mean-PCA, equivalent to risk-premium PCA of Lettau and Pelger, 2020b) and (ii) cross-sectional shape and time-series mean targets combined (XS-shape-TS-mean-PCA).

D.1 Factor weights

This section includes figures that illustrate the in-sample factor weights on the decile portfolios of different models.

For factors that only exploit the cross-sectional shape target (XS-shape-PCA factors), Panels A and B of Figure D.5 show the weights on the extreme-decile portfolios of the first six PCA factors and XS-shape-PCA factors with the penalty on the shape target being $\gamma_{\text{XS}} = 1$, respectively. The horizontal axis in each panel gives the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

In addition, Panels A and B of Figure D.2 give heatmaps for the weights on all decile portfolios of the first six PCA factor and XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$. The horizontal axis depicts the characteristics and the vertical axis the bottom- to top-decile portfolios. Positive (negative) weights are shown in green (red).

Figures D.3 and D.4 show the weights of factors that only exploit the time-series mean target (TS-mean-PCA factors) and alpha target (TS-alpha-PCA factors), respectively. In particular, Panels A and B of Figure D.3 show the weights on the extreme-decile portfolios of the first six PCA factors and TS-alpha-PCA factors with the penalty on the alpha target being $\gamma_{\text{TS}}^{\text{alpha}} = 15$, respectively. Panels A and B of Figure D.4 show the weights on the extreme-decile portfolios of the first six PCA factors and TS-mean-PCA factors with the penalty on the mean target being $\gamma_{\text{TS}}^{\text{mean}} = 15$, respectively.

For factors that exploit both the cross-sectional shape and time-series mean targets (XS-shape-TS-mean-PCA factors), Panels A and B of Figure D.6 show the weights on the extreme-decile portfolios of the first six XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$ and XS-shape-TS-mean-PCA factors with $(\gamma_{\text{XS}}, \gamma_{\text{TS}}^{\text{mean}}) = (1, 15)$, respectively.

For factors that exploit both the cross-sectional macro and time-series mean targets (XS-macro-TS-mean-PCA factors), Panels A and B of Figure D.7 show the weights on the extreme-decile portfolios of the first six XS-macro-PCA factors with $\gamma_{\text{XS}} = 1$ and XS-macro-TS-mean-PCA factors with $(\gamma_{\text{XS}}, \gamma_{\text{TS}}^{\text{mean}}) = (0.5, 15)$, respectively.

D.2 Performance

This section includes figures that illustrate the out-of-sample performance of the TS-mean-PCA and XS-shape-TS-mean models.

Similar to Figure 7 in the main body of the manuscript, Figure D.8 illustrates how the out-of-sample performance of the TS-mean-PCA model varies with the number of factors and the penalty parameter. Panels A, B, C, and D give heatmaps for the out-of-sample Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, pricing error for the FF5-alpha-weighted portfolio, and average unexplained variance across the 370 decile portfolios, respectively. In each panel, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, $\gamma_{\text{TS}}^{\text{mean}}$. Pricing errors are reported in basis points and average unexplained variances in percentage.

To complement Figure 15 in the main body of the manuscript, which illustrates the out-of-sample performance of the five-factor XS-macro-TS-mean-PCA models, Figure D.9 illustrates how the out-of-sample performance of the one- to six-factor models varies with the penalty parameters on the cross-sectional shape and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, and average unexplained variance across the 370 decile portfolios, respectively. Each panel includes six sub-figures that correspond to models with different numbers of factors. In each sub-figure, the vertical axis depicts the penalty on the shape target and the horizontal axis the penalty on the mean target. Root-mean-squared pricing errors are reported in basis points and unexplained variances in percentage.

Figure D.1: PCA and XS-shape-PCA factor weights on extreme-decile portfolios

Panels A and B of this figure show the in-sample weights on the extreme-decile portfolios of the first six PCA factors and XS-shape-PCA factors with $\gamma_{XS} = 1$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) PCA

(B) XS-shape-PCA with $\gamma_{XS} = 1$

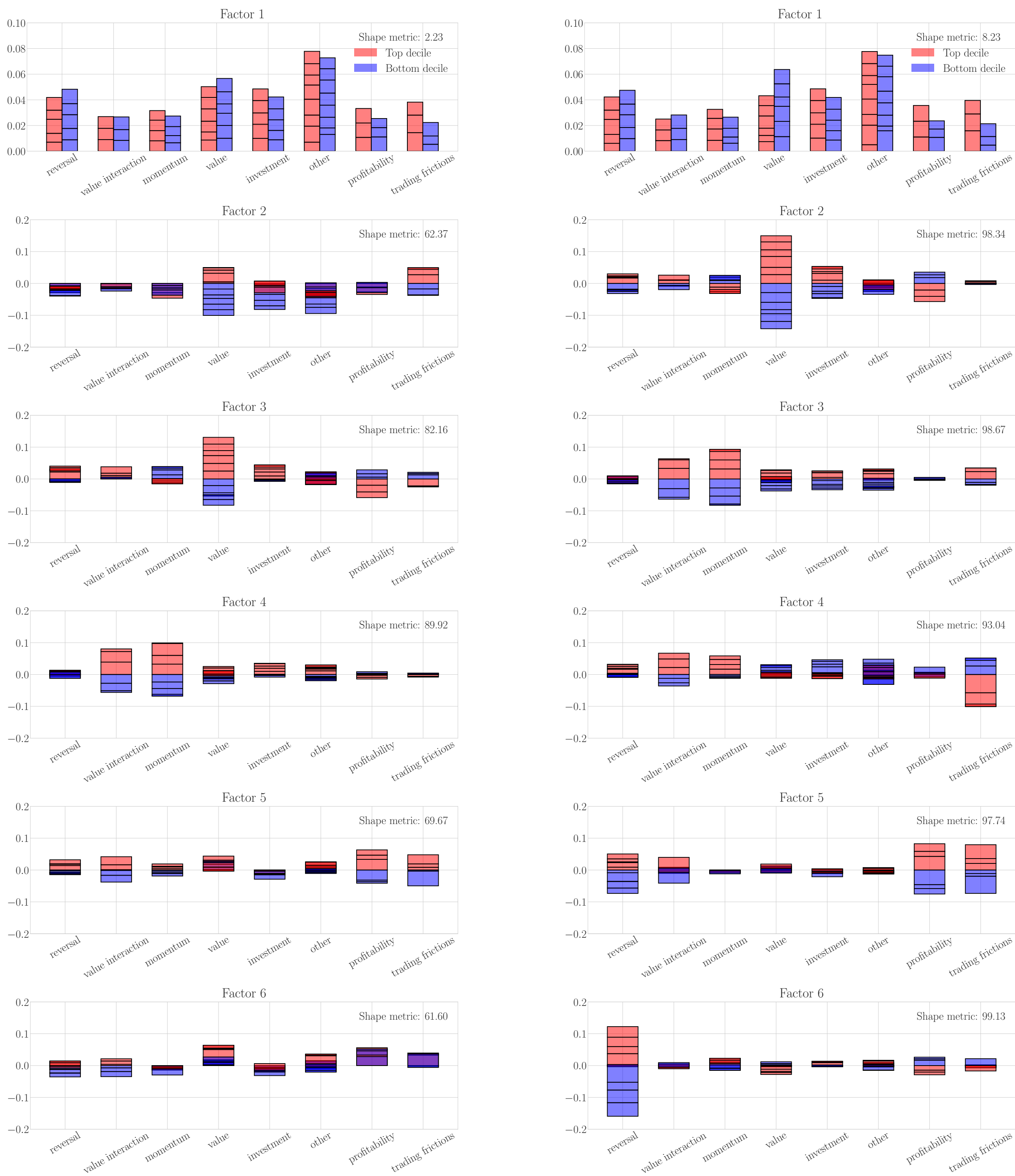


Figure D.2: PCA and XS-shape-PCA weights on all decile portfolios

Panels A and B give heatmaps for the in-sample weights on all decile portfolios of the first six PCA factors and XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$, respectively. The horizontal axis depicts the characteristics and the vertical axis the bottom- to top-decile portfolios. Positive (negative) weights are shown in green (red).

(A) PCA

(B) XS-shape PCA with $\gamma_{\text{XS}} = 1$

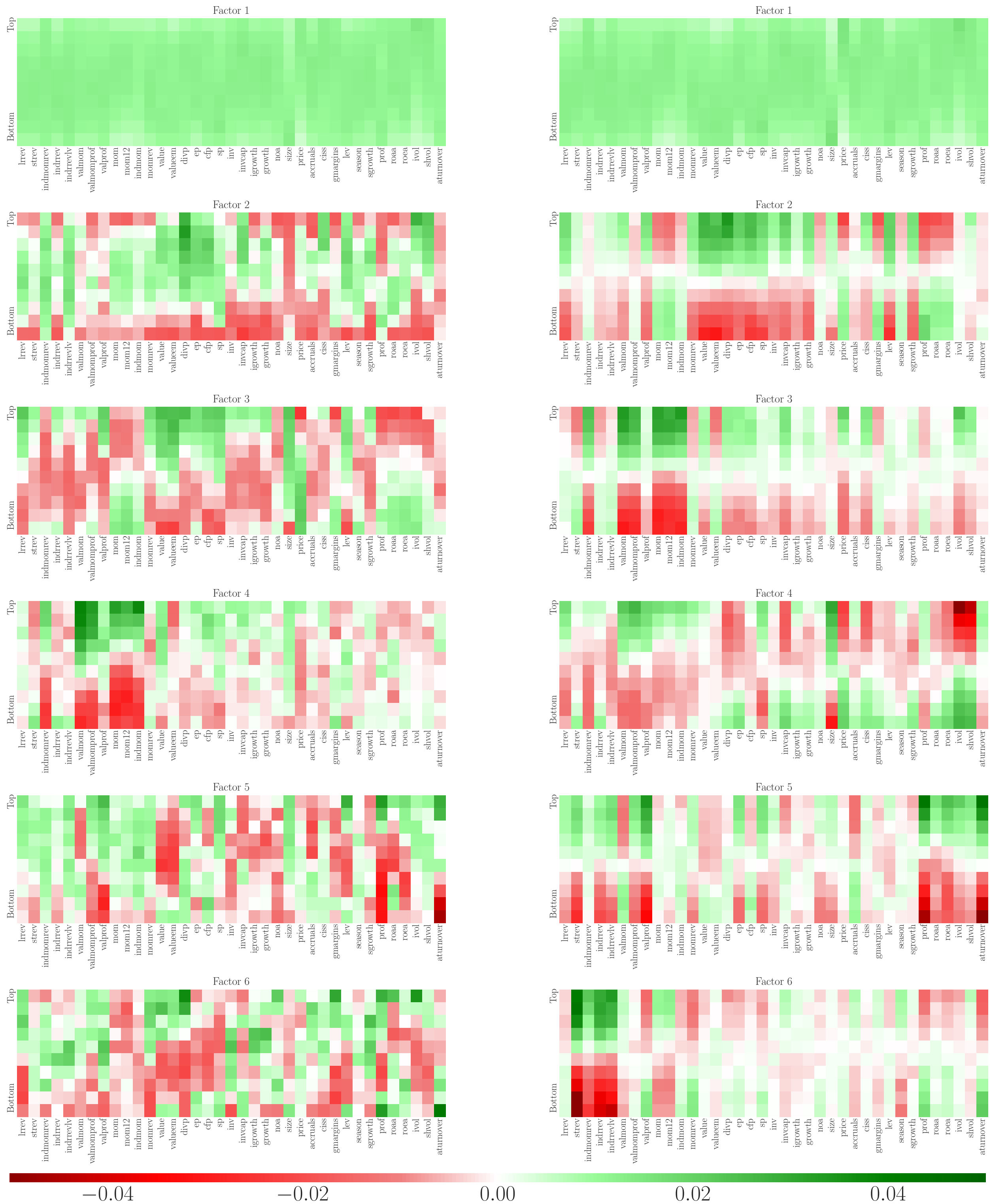


Figure D.3: PCA and TS-alpha-PCA factor weights on extreme-decile portfolios

Panels A and B show the in-sample weights on the extreme-decile portfolios of the first six PCA factors and TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) PCA

(B) TS-alpha-PCA with $\gamma_{\text{TS}}^{\text{alpha}} = 15$

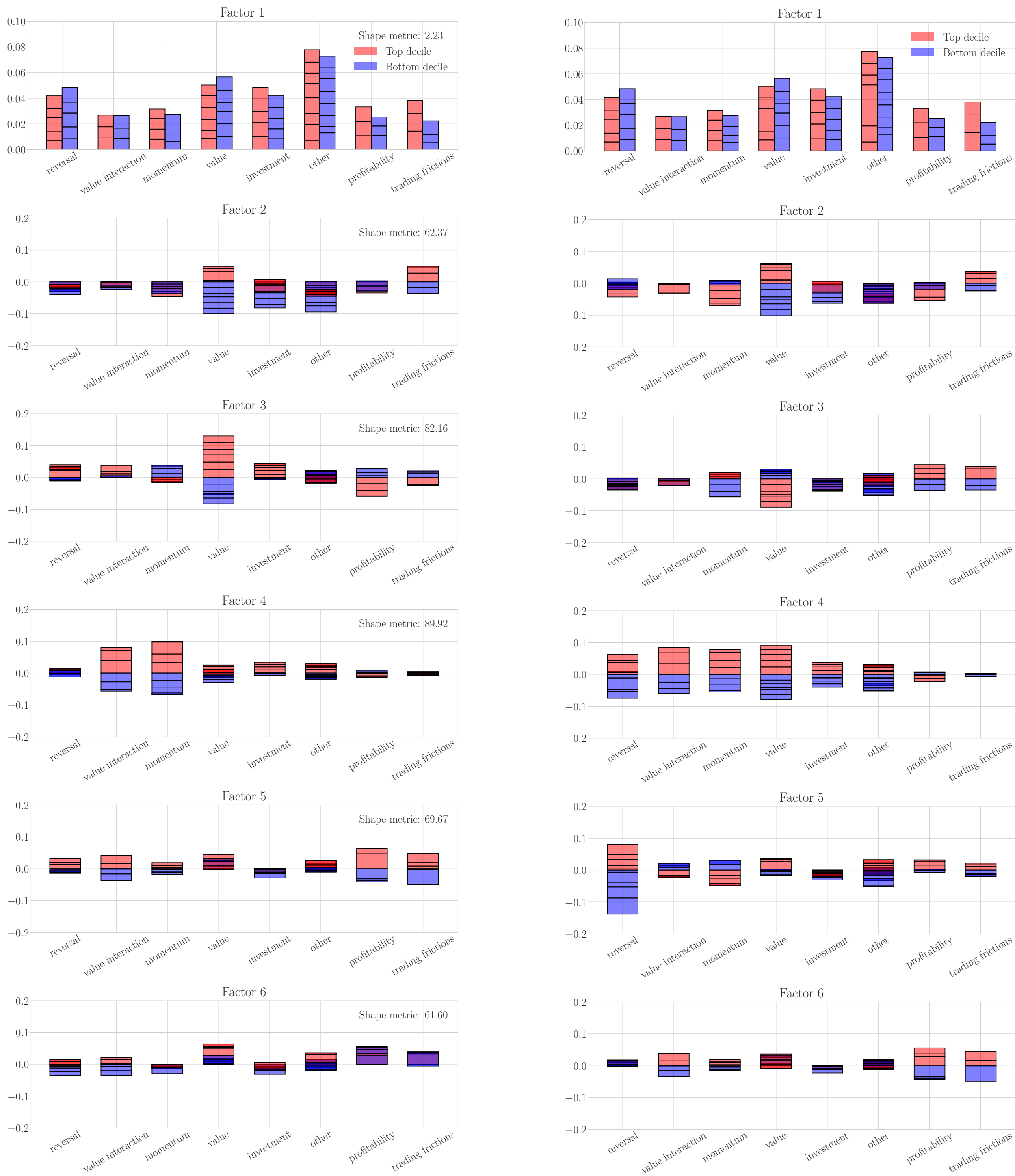
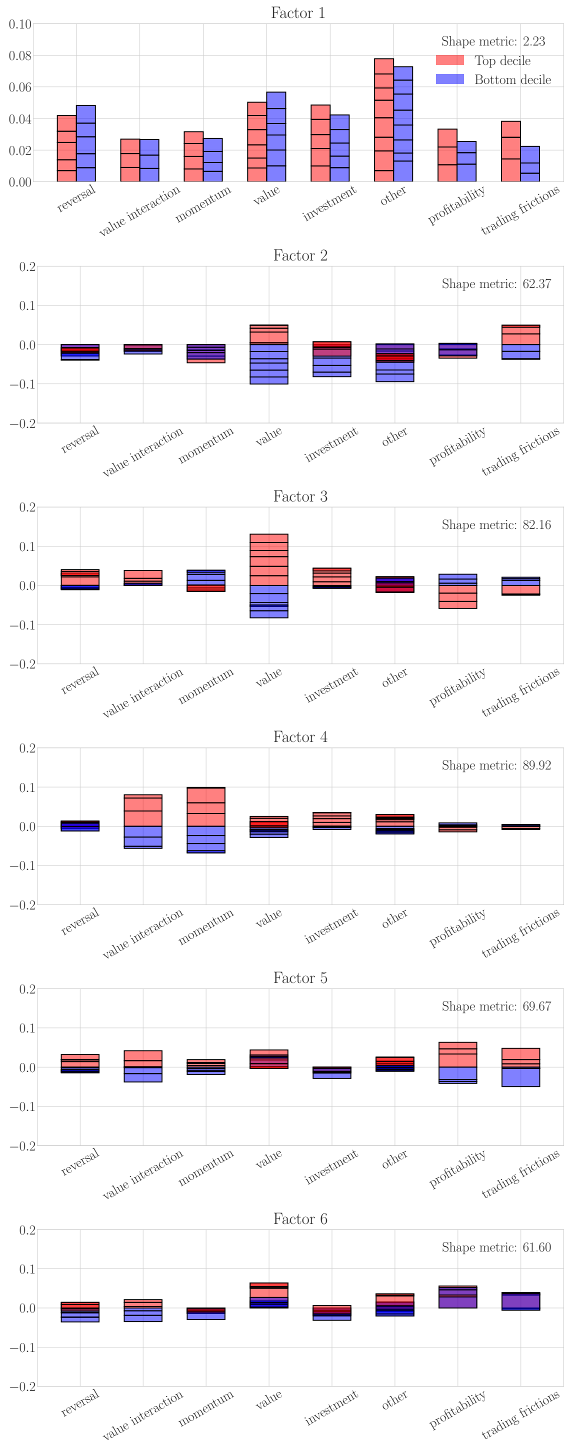


Figure D.4: PCA and TS-mean-PCA factor weights on extreme-decile portfolios

Panels A and B show the in-sample weights on the extreme-decile portfolios of the first six PCA factors and TS-mean-PCA factors with $\gamma_{\text{TS}}^{\text{mean}} = 15$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) PCA



(B) TS-mean-PCA with $\gamma_{\text{TS}}^{\text{mean}} = 15$

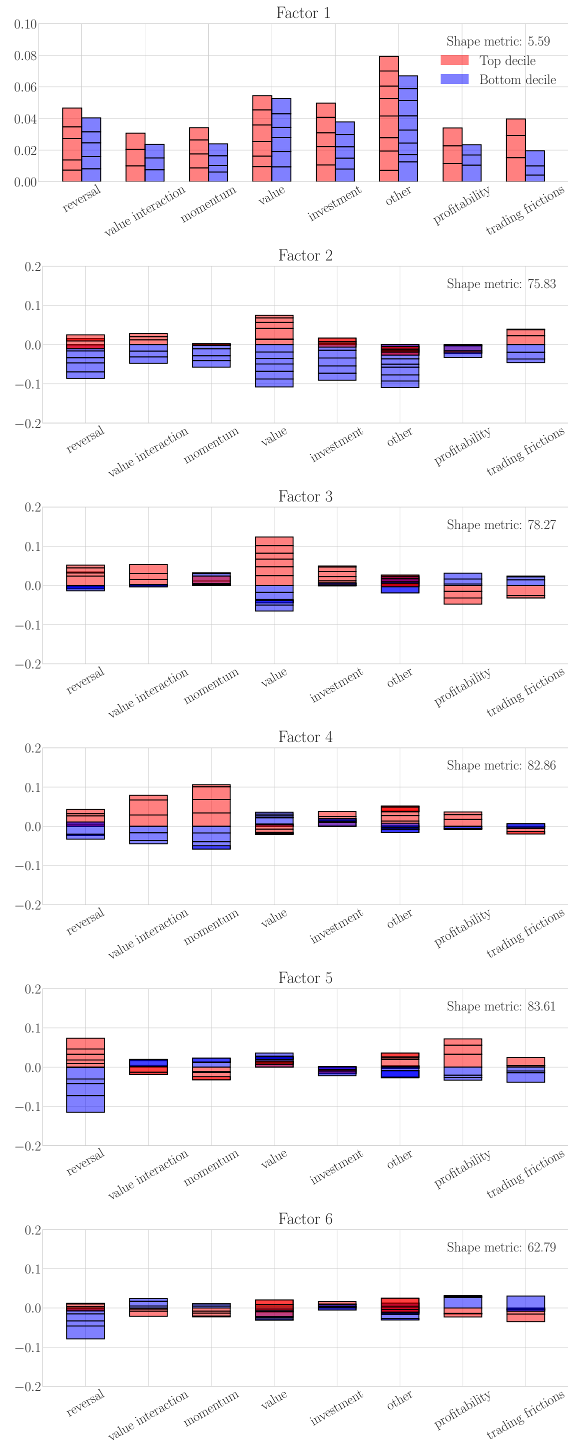


Figure D.5: PCA and XS-shape-PCA factor weights on extreme-decile portfolios

Panels A and B of this figure show the in-sample weights on the extreme-decile portfolios of the first six PCA factors and XS-shape-PCA factors with $\gamma_{XS} = 1$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) PCA

(B) XS-shape-PCA with $\gamma_{XS} = 1$

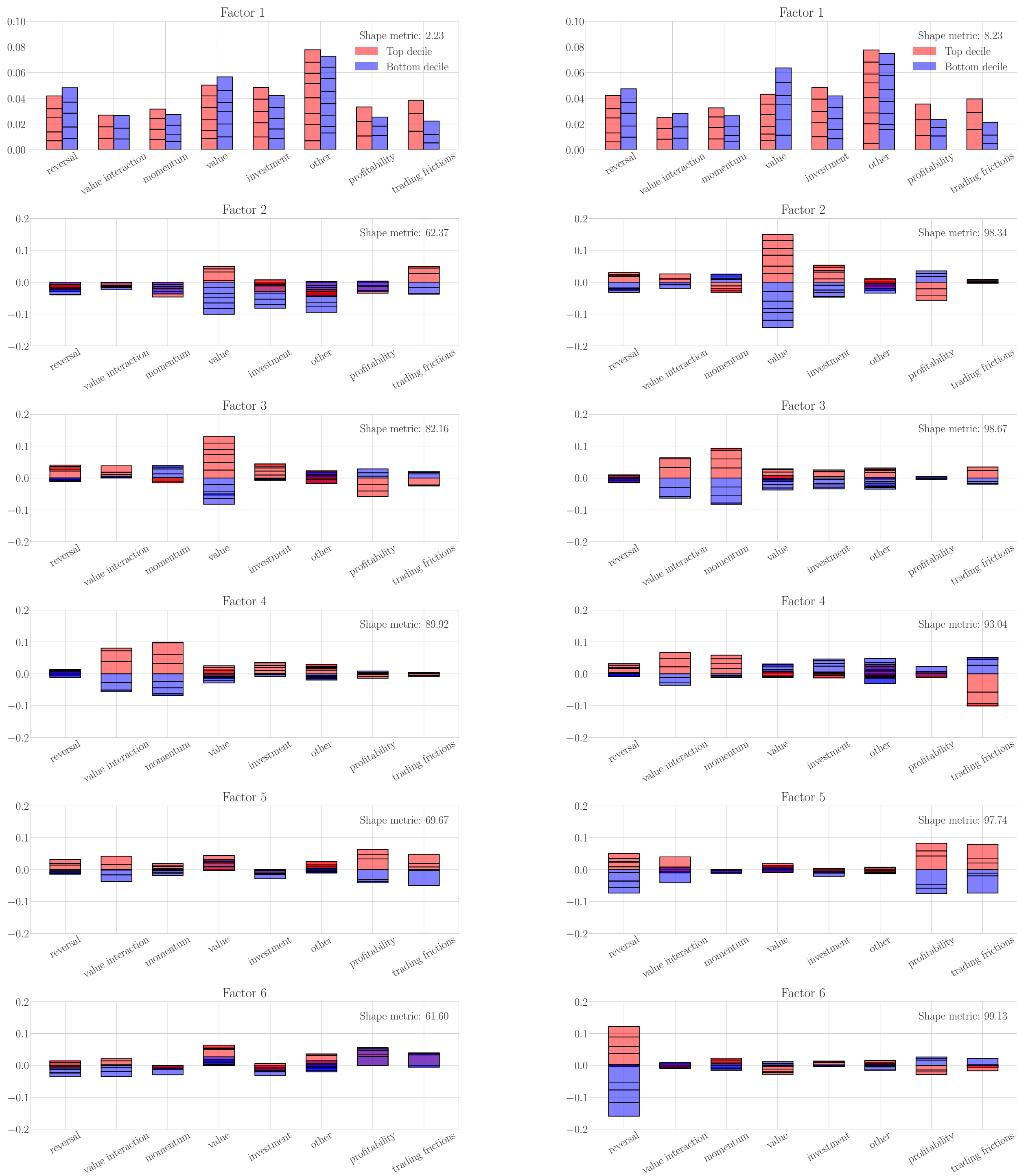


Figure D.6: XS-shape-PCA and XS-shape-TS-mean-PCA factor weights on extreme portfolios

Panels A and B show the in-sample weights on the extreme-decile portfolios of the first six XS-shape-PCA factors with $\gamma_{\text{XS}} = 1$ and XS-shape-TS-mean-PCA factors with $(\gamma_{\text{XS}}, \gamma_{\text{TS}}^{\text{mean}}) = (1, 15)$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) XS-shape-PCA with $\gamma_{\text{XS}} = 1$ (B) $(\gamma_{\text{XS}}, \gamma_{\text{TS}}^{\text{mean}}) = (1, 15)$

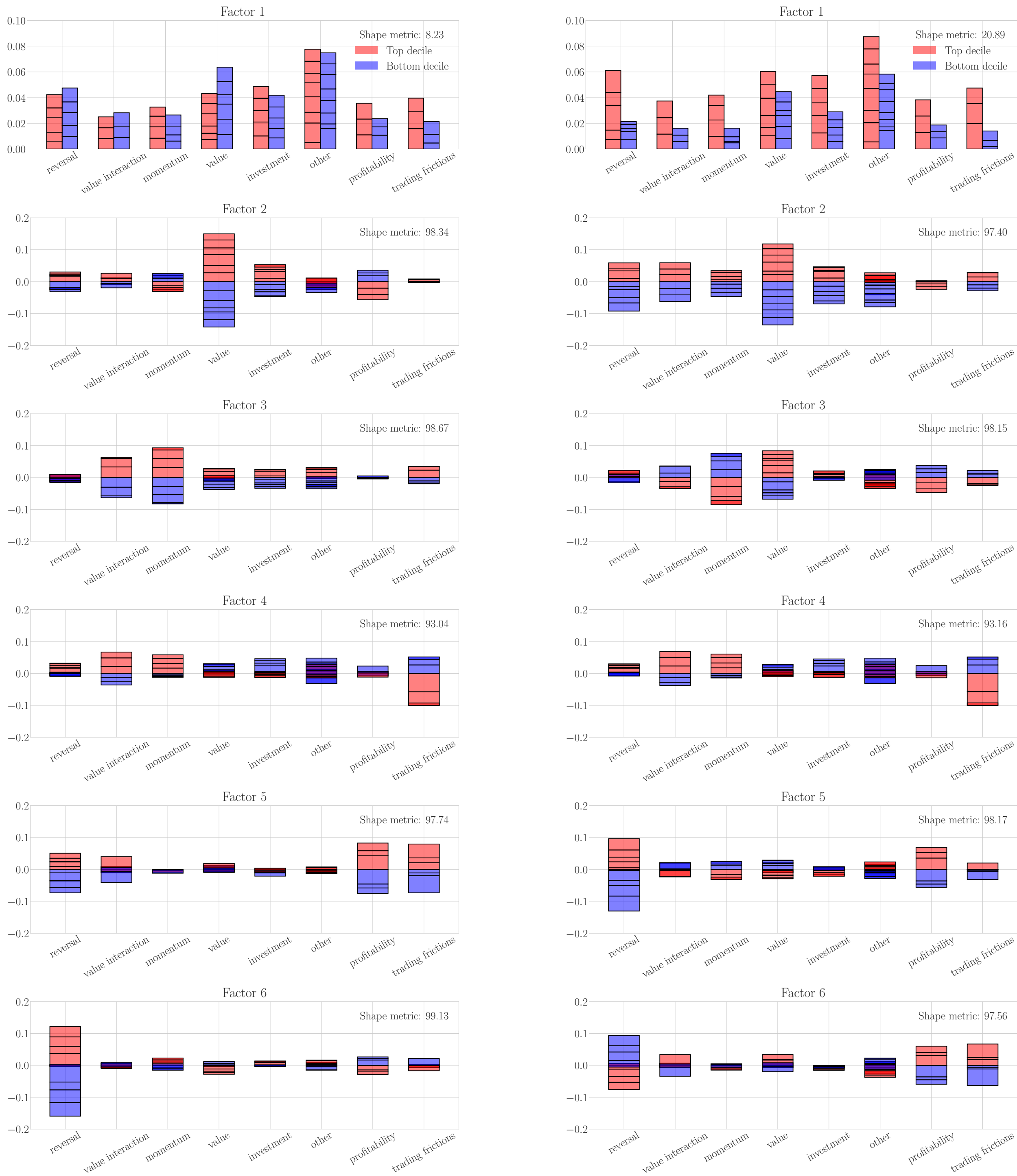


Figure D.7: XS-macro-PCA and XS-macro-TS-mean-PCA factor weights on extreme portfolios

Panels A and B show the in-sample weights on the extreme-decile portfolios of the first six XS-macro-PCA factors with $\gamma_{\mathrm{XS}} = 0.5$ and XS-macro-TS-mean-PCA factors with $(\gamma_{\mathrm{XS}}, \gamma_{\mathrm{TS}}^{\mathrm{mean}}) = (0.5, 15)$, respectively, with the horizontal axis giving the characteristic categories. All factors are normalized to have positive returns. The weights on the top-decile (bottom-decile) portfolios are shown in red (blue). Each bar shows the total weight of a category, and the black lines indicate the contribution of each portfolio in the category.

(A) XS-macro-PCA with $\gamma_{\mathrm{XS}} = 0.5$ (B) $(\gamma_{\mathrm{XS}}, \gamma_{\mathrm{TS}}^{\mathrm{mean}}) = (0.5, 15)$

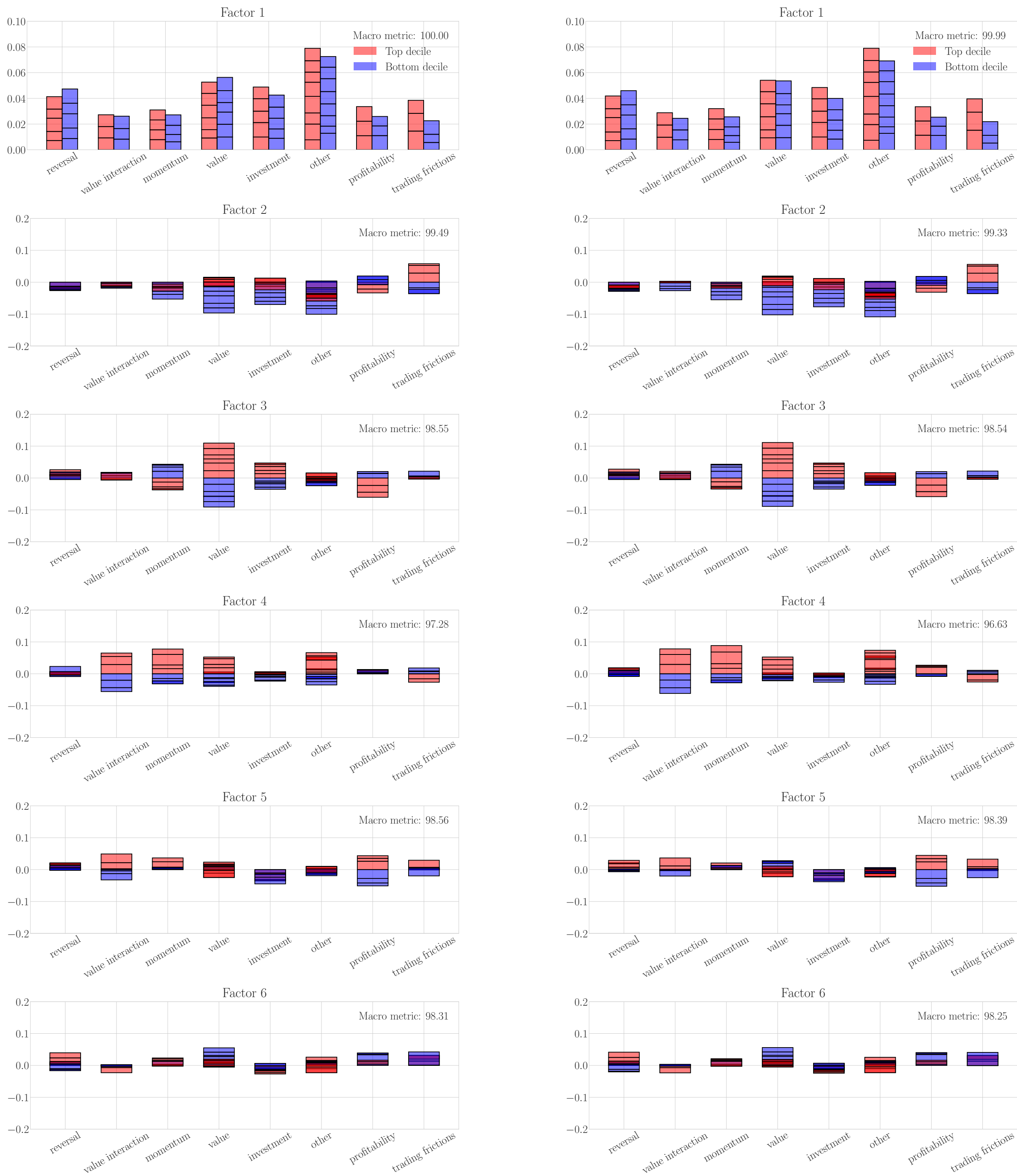


Figure D.8: Performance of time-series mean PCA models

This figure illustrates how the out-of-sample performance of the time-series mean PCA model varies with the number of factors and the penalty parameter. Panel A gives the heatmap for the out-of-sample Sharpe ratio and Panel B for the out-of-sample root-mean-squared pricing error across the 370 decile portfolios. Panel C gives the heatmap for the out-of-sample pricing error for the FF5-alpha-weighted portfolio, which is a portfolio whose weight on each decile is proportional to the decile's alpha with respect to FF5. Panel D gives the heatmap for unexplained variance across the 370 decile portfolios. In each panel, the vertical axis depicts the number of factors in the model and the horizontal axis the penalty parameter, $\gamma_{\text{TS}}^{\text{mean}}$. Pricing errors are reported in basis points and average unexplained variances in percentage.

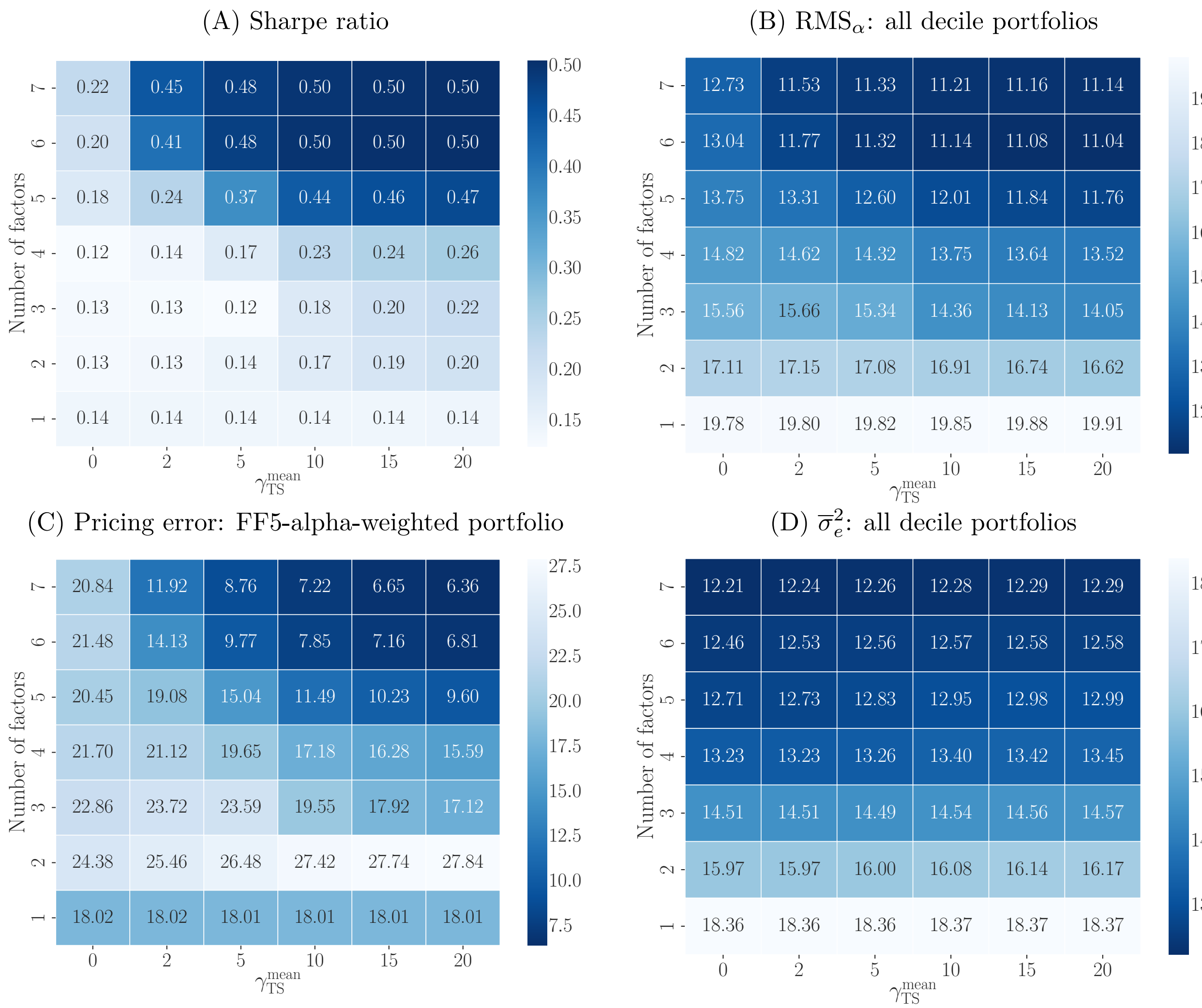


Figure D.9: Performance of XS-shape-TS-mean-PCA models

This figure illustrates how the out-of-sample performance of models varies with the penalty parameters on the cross-sectional shape and time-series mean targets. Panels A, B, C, and D give heatmaps for the Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error across the 37 top-minus-bottom portfolios, and average unexplained variance across the 370 decile portfolios, respectively. Each panel includes six sub-figures that correspond to models with different numbers of factors. In each sub-figure, the vertical axis depicts the penalty on the shape target and the horizontal axis the penalty on the mean target. Root-mean-squared pricing errors are reported in basis points and unexplained variances in percentage.

(A) Sharpe ratio

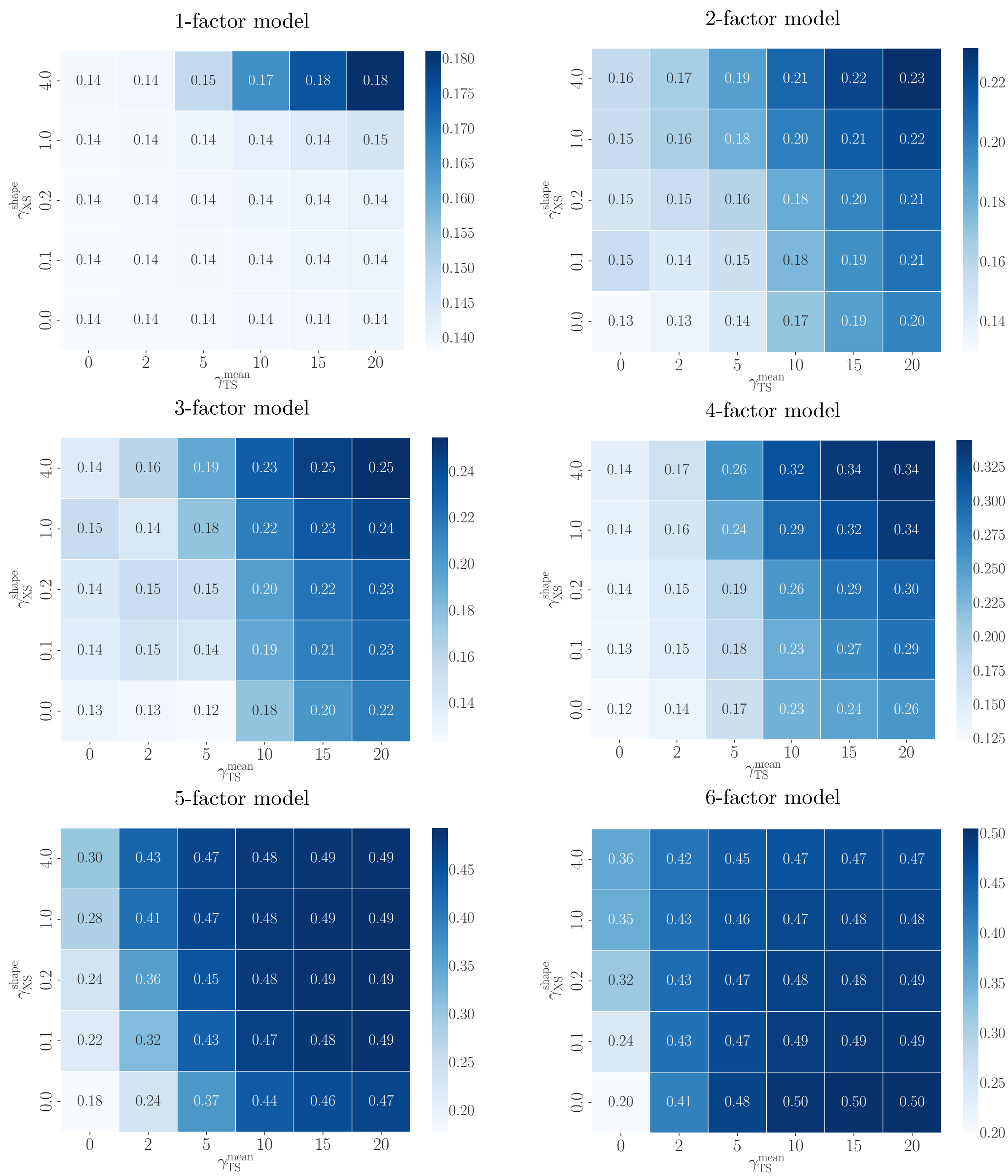


Figure D.9 continued

(B) RMS_α: all decile portfolios

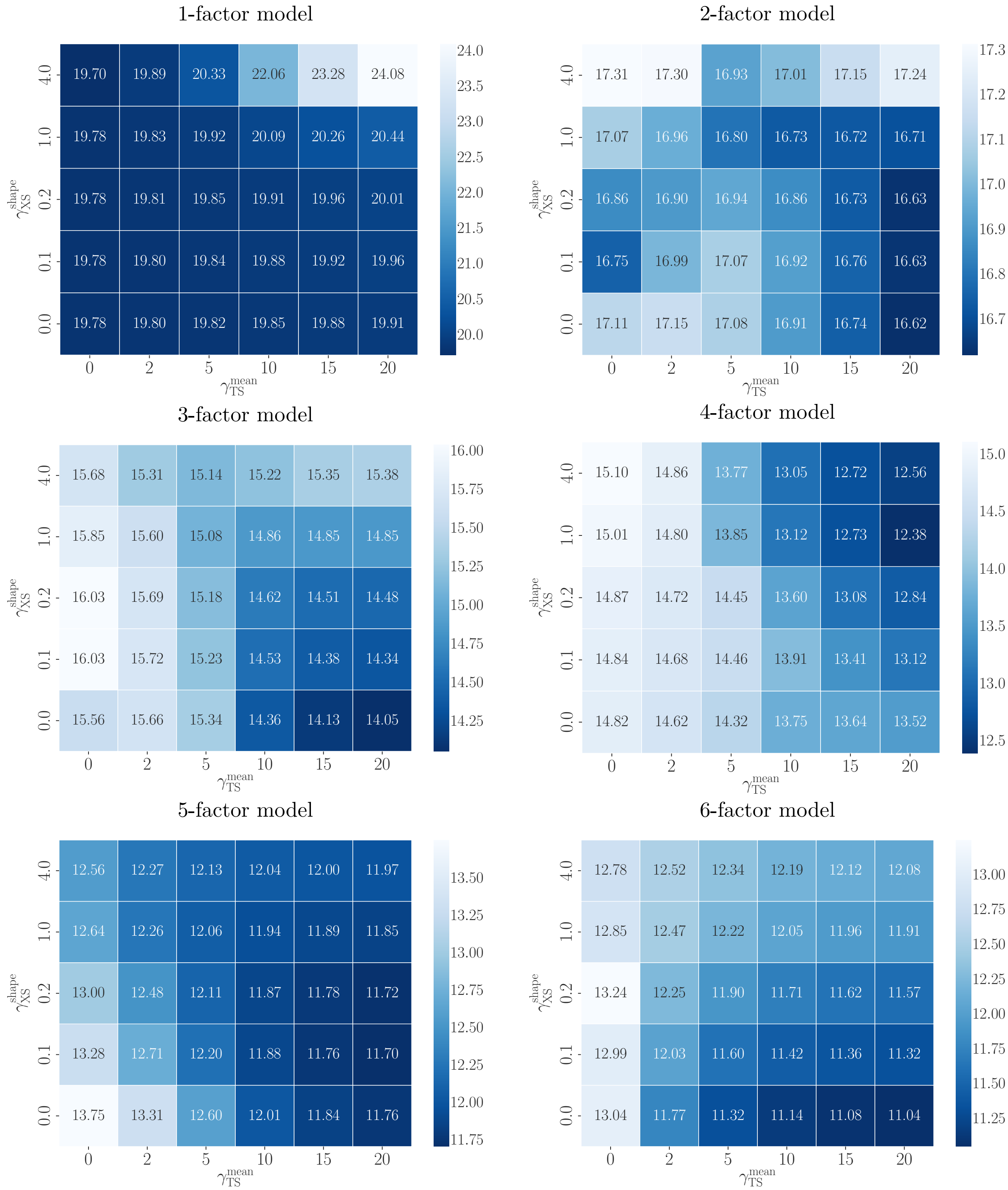
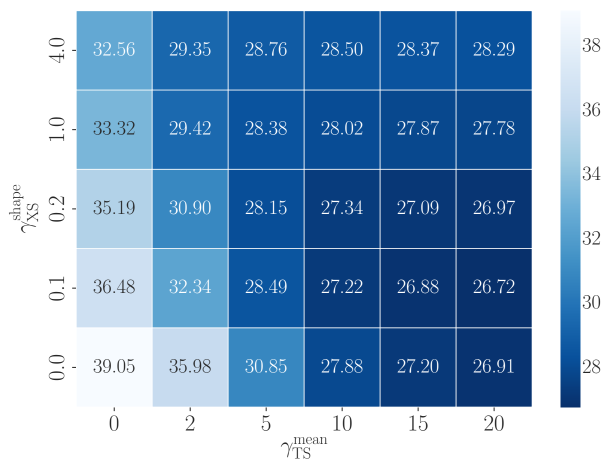


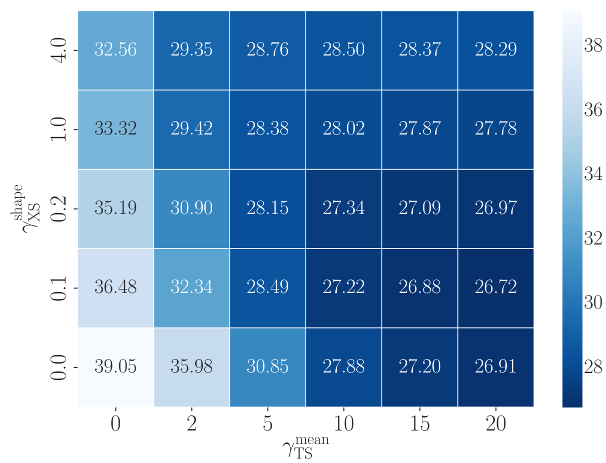
Figure D.9 continued

(C) RMS$_\alpha$: top-minus-bottom portfolios

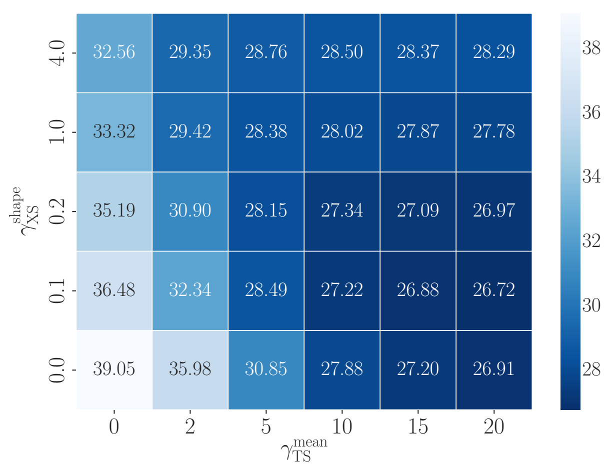
1-factor model



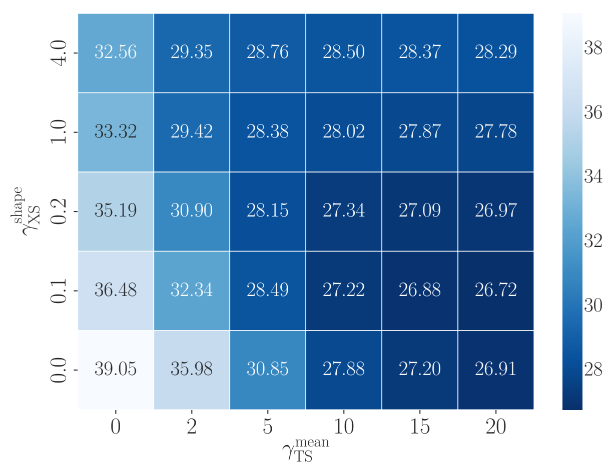
2-factor model



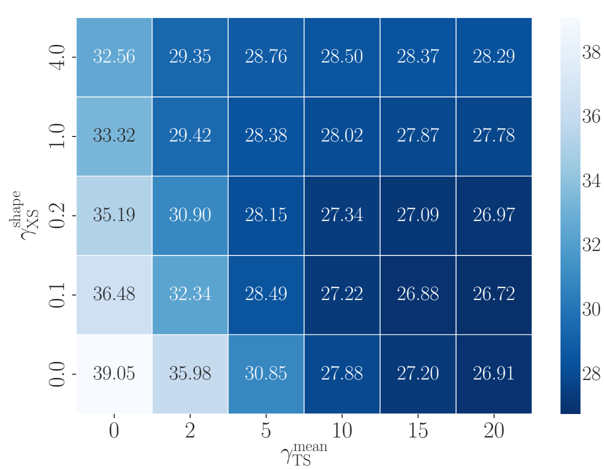
3-factor model



4-factor model



5-factor model



6-factor model

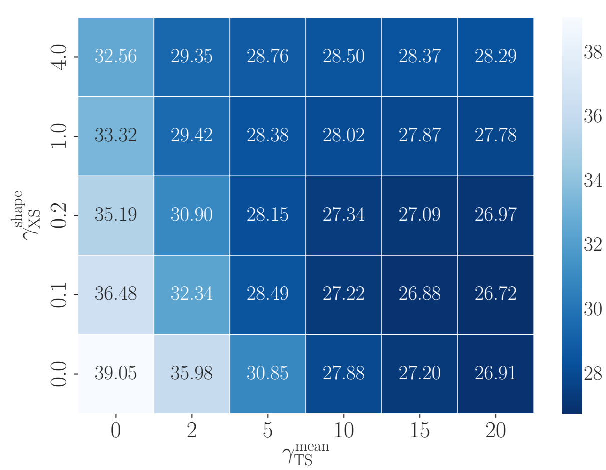


Figure D.9 continued

(D) $\overline{\sigma}_e^2$: all decile portfolios