Asset-Pricing Factors with Economic Targets*

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

August 27, 2025

Abstract

We develop a general framework that integrates economic insight into the estimation of latent factors to improve asset-pricing models. Extending principal component analysis, our method incorporates economically motivated cross-sectional and time-series moment conditions to identify even weak factors. Cross-sectional targets can enforce economic structure, such as monotonicity in factor loadings or alignment with macroeconomic shocks. Time-series targets rewards factors that better explain expected returns or reduce pricing errors relative to benchmark models. In an extensive empirical study, we find that these economically guided risk factors better span the pricing kernel, yielding significantly higher Sharpe ratios and lower mispricing than conventional approaches.

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

*We are grateful for comments from Zareei Abalfazl (discussant), Martijn Boons, Joao Cocco, James Dow, Elise Gourier, Raymond Kan (discussant), Nathan Lassance, Alberto Martin-Utrera, Shane Miller (discussant), Andreas Neuhierl (discussant), Jean Pauphilet, Anna Pavlova, Seth Pruitt, Francesco Roccazzella, Mirela Sandulescu (discussant), Stephen Schaefer, Raman Uppal, Frédéric Vrins, Dacheng Xiu, Paolo Zaffaroni, Xinran Zhang (discussant) and seminar and conference participants at London Business School, and Université Catholique de Louvain, the 16th International Conference on Computational and Financial Econometrics, 2022 INFORMS Annual Meeting, 2023 CICF Conference, 2023 EFA Annual Meeting, 2023 Financial Econometrics Conference at Lancaster University to mark Stephen Taylor's Retirement, 2023 NFA Annual Meeting, 2023 Oxford Man Institute Conference on Machine Learning and Financial Econometrics, 2023 SoFiE Conference, 2023 WFA Annual Meeting, 2024 AFA Annual Meeting, 2024 Hong Kong Conference for Fintech, AI, and Big Data in Business, and 2024 Asset Pricing Workshop at Alliance Sorbonne.

†London Business School and CEPR, sbryzgalova@london.edu

‡London Business School, avmiguel@london.edu

§The Chinese University of Hong Kong (CUHK) Business School, sicongli@cuhk.edu.hk

¶Stanford University, Department of Management Science & Engineering, mpelger@stanford.edu.

1 Introduction

A fundamental insight of asset pricing is that expected returns should be explained by asset exposure to systematic risk factors, and thus, a critical research challenge in the last few decades has been to identify appropriate risk factors. This research has led to a "factor zoo," with more than 300 candidate factors proposed by researchers to explain the cross section (Cochrane, 2011; Harvey, Liu, and Zhu, 2016). In an attempt to bring order to the factor chaos and find a parsimonious factor model that explains the cross-section, researchers have resorted to powerful statistical methods to reduce dimensionality, with Principal Component Analysis (PCA) and its variations playing a prominent role. However, statistical factor analysis based on PCA extracts factors that capture comovement but neglects other return moments and any economic structure, and as a result cannot detect weaker factors that are relevant for asset pricing. In this paper, we show how to overcome this challenge by incorporating economic information into statistical learning.

We propose a general framework to encode economic information into latent-factor modeling. Our method combines the flexibility of data-driven estimation with the insights from structural and reduced-form asset-pricing models. Instead of selecting an ad-hoc set of factors, we let the data speak, but use economic knowledge to nudge the factors to have desirable economic properties and detect weak signals. Our estimator generalizes PCA by including a reward for general cross-sectional (XS) and time-series (TS) moment targets. Economically motivated cross-sectional targets can capture monotonicity constraints on the loadings of factors or their correlation with fundamental macroeconomic innovations. Time-series targets can reward explaining expected returns or reducing mispricing relative to a benchmark reduced-form model. In an extensive empirical study, 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.

Our approach, which we term cross-sectional time-series target PCA (XS-TS-target-PCA), offers a disciplined framework to examine which cross-sectional and time-series target information is useful for asset pricing. While our framework allows for a general formulation of the cross-sectional targets, we focus on two particular instances. First, we show how to impose economically motivated shape restrictions on risk-factor loadings.[1] Most portfolios of stocks sorted by firm characteristics exhibit a monotonic pattern in expected returns, which routinely leads to the construction of long-short factors that capture a significant spread in asset returns.[2] Motivated by this observation, we nudge our latent factors so that their pricing span reflects the monotonicity in the expected returns of characteristic-sorted portfolios. This introduces economic structure based on risk-premia information and tilts the latent factors to have an interpretable long-short structure. Second, we estimate latent factors that are correlated with fundamental macroeconomic innovations. Tradable factors are often thought to reflect systematic sources of risk originated by nontradable (and em-

[1] This approach is inspired by the macro literature that uses shape restrictions to identify structural shocks, pioneered by the long-run restrictions of Blanchard and Quah (1989) and the sign restrictions of Uhlig (2005).

[2] Examples include Patton and Timmermann (2010); Hou, Xue, and Zhang (2018); Jensen, Kelly, and Pedersen (2023).

pirically poorly measured) state variables (macroeconomic innovations or shocks to intermediary risk capacity). We incorporate this insight by explicitly constraining the pricing span of our latent factors to include that of exogenous sources of risk, captured by their covariance with asset returns. Thus, we nudge our factors to directly reflect the asset-pricing implications of state variables that have been identified by economic theory.

We also exploit information in asset risk premia by incorporating time-series moments as economic targets. We again allow for a general formulation of the targets, and focus on two particular cases. First, risk factors should not only capture the source of systematic time-series variation in asset returns, but, consistent with Arbitrage Pricing Theory (APT), also carry risk premia that explain cross-sectional differences in expected returns. The risk-premium PCA of Lettau and Pelger (2020b) is designed to achieve this 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 risk-premium PCA as a particular type of time-series target. Second, we propose the *alpha PCA* approach, which nudges latent factors to span 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 Fama and French (2015) five-factor model (FF5), and reward statistical risk factors that capture the expected returns that the benchmark model *fails to explain*. This target can improve upon a simple risk-premium target because cross sections can be dominated by "easy" factors. The alpha target nudges the estimator to explore hard-to-price assets and put more weight on the weak factors that explain the returns of the alpha portfolios.

The economic targets provide information that helps to detect weak factors that conventional approaches miss. A key element in the estimation of latent factors is factor strength. Intuitively, the strength of a factor in a given panel depends on how much variation is explained by the factor. A factor that affects only a small portion of the assets or explains little variation is a weak factor. PCA estimation applied to a sample covariance matrix can consistently estimate only latent factors that are sufficiently strong. Thus, PCA may fail to detect weak factors, even when they carry large risk premia and constitute an important component of the pricing kernel. Our estimator applies PCA to a modified second-moment matrix that incorporates the target information, and hence is trivial to implement. Importantly, our method increases the strength of weaker factors. As a result, economic targets can provide the information that is necessary to detect and consistently estimate factors that conventional PCA misses.

We empirically evaluate how different economic targets affect asset-pricing models using a large dimensional panel of 370 decile portfolios constructed from single sorts on 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 that of the model based on the combination of two types of targets. We highlight five key insights from our empirical analysis. First, we find that imposing cross-sectional shape restrictions (consistent with the monotonic patterns in expected returns of characteristic-sorted portfolios) significantly improves the empirical performance

of latent asset-pricing factors, which achieve higher out-of-sample Sharpe ratios and lower pricing errors compared to standard PCA factors. Moreover, the resulting XS-shape-PCA factors have a stronger long-short structure in characteristics compared to conventional PCA factors, particularly for characteristics in the reversal, value-interaction, profitability, and trading-friction categories. Second, incorporating either of the time-series targets in factor estimation (mean return or alpha) also leads to a significant improvement in 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 a broad range of characteristic-sorted portfolios, the TS-alpha-PCA factors better capture the average returns of portfolios that are notoriously challenging to price for the FF5 model such as those in the reversal category.

Our third empirical observation is that there are benefits from combining the cross-sectional and time-series targets. In particular, imposing both shape restrictions on the factor-loading pattern as well as APT-driven time-series pricing restrictions leads to risk factors that strongly reflect the underlying long-short pattern of characteristic-sorted portfolio returns, while also achieving high Sharpe ratios and low pricing errors. In other words, imposing a cross-sectional shape target on top of a time-series mean or alpha target not only leads to an improvement in asset-pricing performance, but also enhances the interpretability of the resulting latent factor models. Our fourth finding is that although imposing a cross-sectional macro target can help to identify factors that are better at pricing the mimicking portfolios of macroeconomic variables, it also leads to factors that are worse at pricing a general cross section of assets. Thus, there is a tradeoff between pricing the macro mimicking portfolios and a broad cross-section of assets. Finally, 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 method is grounded in formal asymptotic theory. The econometric theory and statistical arguments for the cross-sectional target are based on Duan, Pelger, and Xiong (2023) and for the time-series target on Lettau and Pelger (2020a). While Duan et al. (2023) 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. Lettau and Pelger (2020a) study the properties of a mean target, but similar arguments apply to an alpha or other time-series targets.

Approximate factor models are one of the most popular frameworks used to explain a wide cross-section of asset returns (Ross, 1976; Chamberlain and Rothschild, 1983). While the conditions and methods for inference on strong factors are well established (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 reduce estimation uncertainty and improve factor identification.

Our paper contributes to the literature on high-dimensional factor modeling and our approach is complementary to many PCA-based approaches. In particular, our method can be combined with existing approaches by simply replacing the conventional PCA factors with our XS-TS-target-PCA factors in order to incorporate additional economic information. Examples of complementary methods include the instrumented PCA of Kelly, Pruitt, and Su (2019) and the closely related projected PCA of Fan, Liao, and Wang (2016) and regressed PCA of Chen, Roussanov, and Wang (2023) and Chen, Roussanov, Wang, and Zou (2024b), the robust stochastic discount factor (SDF) of Kozak, Nagel, and Santosh (2020), the high-frequency PCA of Pelger (2020), the three-pass model of Giglio and Xiu (2021), and the local PCA of Zaffaroni (2022). While these papers assume a strong factor model, our approach allows one to generalize their methods to the case with weaker signals.

Our paper also contributes to the emerging literature that estimates weak factors for asset pricing. The closely related paper by Lettau and Pelger (2020b) introduces the APT-driven requirement for risk factors to span expected returns of the portfolios (in addition to their time-series variation). The objective is to find weak factors that explain the cross-section of expected returns. Our general framework nests their risk-premium PCA as a particular case of time-series target, which we include in our empirical study. Our approach is more general because it not only allows one to include different types of time-series targets, but also impose additional cross-sectional restrictions on the shape and structure of latent factors. A complementary approach is proposed by Giglio, Xiu, and Zhang (2022). Their supervised PCA—and similarly the scaled PCA of Huang, Jiang, Li, Tong, and Zhou (2021)—estimates latent factors that have the highest correlation with a time-series target. To do this, they apply PCA to the subset of test assets that are most highly correlated with the target. Their objective is to estimate the risk premia of the time-series target by constructing traded target-mimicking factor portfolios. Thus, their approach can detect weak factors that are highly correlated with the target, while our approach can detect weak factors that satisfy economic constraints and are useful for pricing the overall cross section. The two approaches are complementary, as they pursue related but distinct goals, and can be combined by replacing the conventional PCA factors in supervised PCA with the XS-TS-target-PCA factors to incorporate additional economic information. Dello Preite, Uppal, Zaffaroni, and Zviadadze (2022) propose an alternative approach based on a constrained maximum likelihood SDF estimation that includes a no-arbitrage bound, which is a specific type of economic constraint. They show that unsystematic risk is priced, and that the unsystematic SDF component maps into weak cross-sectional factors.

More broadly, our work is related to the literature that includes economically motivated structure or regularization in flexible data-driven methods. Chen, Pelger, and Zhu (2024a) estimate the conditional SDF with deep neural networks using a no-arbitrage objective function, while Bryzgalova, Pelger, and Zhu (2019) use decision trees for a robust SDF recovery. Filipovic, Pelger,

and Ye (2022a,b) estimate arbitrage-free yield curves and term-structure factors with a smoothness penalty, which is equivalent to a penalty on excessive payoffs. Bryzgalova, Huang, and Julliard (2022) impose economic restrictions to recover the underlying SDF. Bali, Goyal, Huang, Jiang, and Wen (2022) predict bond returns with machine learning by including information from a structural economic model.[3]

2 Factor models with economic targets

2.1 Factor model

Assume that excess returns follow a standard approximate factor model, with dynamics 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:

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

or, in matrix notation,

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

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 residuals are uncorrelated, the covariance matrix of asset returns can be decomposed into systematic and idiosyncratic parts:

$$\Sigma_R = \beta \Sigma_F \beta^\top + \Sigma_e. \tag{1}$$

If factors are sufficiently strong, that is, they explain enough time-series variation in the panel, we can estimate the factors F via 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

[3] More generally, our paper is complementary to the growing literature on return prediction with machine learning, which has shown the benefits of using regularized flexible methods. In their seminal work, Gu, Kelly, and Xiu (2020) compare machine-learning methods for predicting the panel of individual U.S. stock returns. Freyberger, Neuhierl, and Weber (2020) use a nonlinear adaptive group LASSO to predict stock returns. Bianchi, Buchner, and Tamoni (2022) and Li and Rossi (2020) use machine learning to predict bond returns and mutual-fund returns, respectively. A related stream extends cross-sectional machine-learning prediction to higher moments and residuals, which are used for investment decisions. For instance, Li and Tang (2022) use machine learning to predict risk measures and estimate conditional volatilities and Kaniel, Lin, Pelger, and Van Nieuwerburgh (2023) and DeMiguel, Gil-Bazo, Nogales, and Santos (2023) predict the skill of mutual-fund managers.

factors and the variance of the factors is sufficiently large. Formally, a conventional PCA estimator minimizes the following objective function:

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

where $\bar{R}$ and $\bar{F}$ are the sample mean of the returns and factors and $\|\cdot\|_F$ is the Frobenius norm. Under the standard normalization that the loadings are orthonormal $\hat{\beta}_{\text{PCA}}^\top \hat{\beta}_{\text{PCA}}/N = I_K$, the PCA estimator follows from the 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 $\hat{F}_{\text{PCA}} = R\hat{\beta}_{\text{PCA}}(\hat{\beta}_{\text{PCA}}^\top \hat{\beta}_{\text{PCA}})^{-1}$. In other words, latent factors are linear combinations (portfolios) of the original returns R, with portfolio weights $\hat{\omega}_{\text{PCA}} \in \mathbb{R}^{N \times K}$:

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

As long as factors are uncorrelated, the factor weights and conventional OLS loading estimators coincide, 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. PCA aims to find latent factors that explain as much time-series variation as possible for a given set of basis assets, and thus, it only leverages information in the second moment (covariance matrix), and ignores all other information, such as other moments (including the mean) and potential economic restrictions. Furthermore, it assigns the same weight to 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 small subset of the cross-section or have a low variance might not be detectable with PCA.

A promising solution is 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 their exposure to factors and the factor means, that is, $E[R] = \beta E[F]$. 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 can leverage information from both cross-section and time-series moments of returns and additional data (e.g., macroeconomic innovations) in recovery of the asset-pricing factors, which nests 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 (whose 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 a benchmark reduced-form model. By 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 incorporate cross-sectional targets into latent-factor estimation. The cross-sectional target PCA (XS-target-PCA) approach adds a penalty to estimate factors that align with an economically motivated cross-sectional structure. In the general case, we model the cross-sectional target as a matrix $\Lambda^{\text{XS}} \in \mathbb{R}^{N\times L}$, and our method pushes the matrix of factor loadings β to be correlated with the target. Formally, this pushes the 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.

To provide some intuition behind the estimator, we assume that the target matrix Λ^{XS} is normalized to be orthonormal, that is $(\Lambda^{\text{XS}})^\top \Lambda^{\text{XS}}/N = I_L$. This normalization is without loss of generality and does not constraint our approach. 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 the returns of both the basis assets R and the target portfolios R^{XS}. Hence, the estimator solves the following optimization problem

$$\underset{F,\beta}{\text{argmin}} \frac{1}{NT}\sum_{n=1}^{N}\sum_{t=1}^{T}\left(R_{t,n} - F_t\beta_n^\top\right)^2 + \gamma_{\text{XS}}\frac{1}{LT}\sum_{l=1}^{L}\sum_{t=1}^{T}\left(R^{\text{XS}}_{t,l} - F_t(\beta^{\text{XS}}_l)^\top\right)^2, \tag{3}$$

where $(\beta^{\text{XS}})^\top = \beta^\top \Lambda^{\text{XS}}/N$ is the beta of the target portfolios and γ_{XS} is the cross-sectional target penalty. In the special case of $\gamma_{\text{XS}} = 0$, we recover a conventional PCA estimator.[4] Increasing the penalty γ_{XS} pushes the latent factors to better explain the returns of the managed 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* $P_\Lambda^{\text{XS}} \in \mathbb{R}^{N\times N}$

$$P_\Lambda^{\text{XS}} = \Lambda^{\text{XS}} \left((\Lambda^{\text{XS}})^\top \Lambda^{\text{XS}} \right)^{-1} (\Lambda^{\text{XS}})^\top, \tag{4}$$

which projects N dimensional vectors onto the column space of Λ^{XS}. Thus, the XS-target-PCA approach estimates factors using the following objective:

$$\text{XS-target-PCA:} \quad \hat{F}_{\text{XS}}, \hat{\beta}_{\text{XS}} = \underset{F,\beta}{\text{argmin}} \underbrace{\frac{1}{NT} \left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{variation of returns}}} + \gamma_{\text{XS}} \underbrace{\frac{1}{LT} \left\| RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}} \right\|_F^2}_{\substack{\textbf{XS target}\text{: unexplained variation} \\ \text{of } P_\Lambda\text{-projected returns}}}. \tag{5}$$

Note that problems (3) and (5) are equivalent when 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 of returns R, then setting γ_{XS} sufficiently large allows one to detect these weak factors. Note that the dimensions of the panel of basis assets and target assets can be very different, but this 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 (5) 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*

$$\frac{1}{NT} R \left(I_N + \gamma_{XS} \frac{N}{L} P_\Lambda^{XS} \right) R^\top. \tag{6}$$

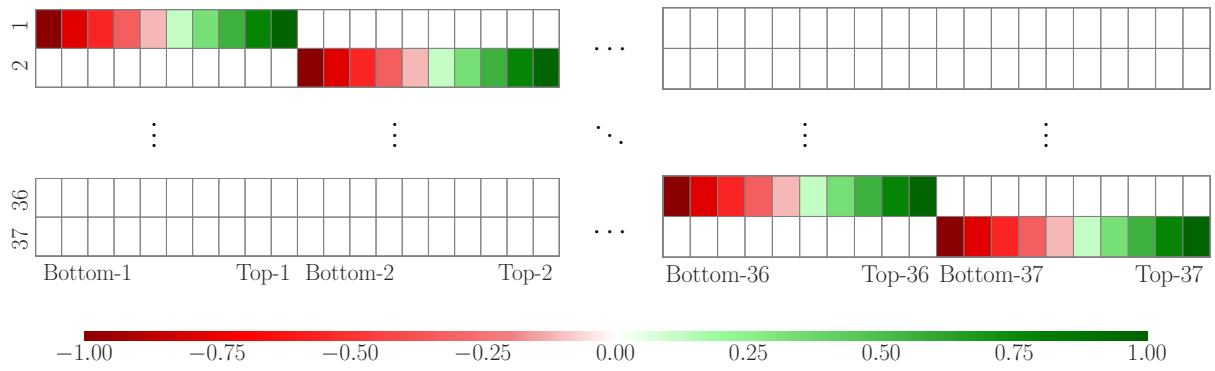
Specifically, the estimated factor returns, $\hat{F}_{XS}$, are the first K eigenvectors of matrix (6), and the estimated factor loadings are $\hat{\beta}_{XS} = R^\top \hat{F}_{XS}$.

In our empirical analysis we consider two types of cross-sectional targets. The first target imposes economically motivated shape restrictions on the structure of the loadings. Most asset-

[4] Note that the PCA problem in (2) differs from problem (3) for the case with $\gamma_{\text{XS}} = 0$ because while problem (2) is formulated in terms of the *demeaned* returns, problem (3) is defined in terms of uncentered returns. However, the estimators obtained using problem (2) and problem (3) for the case with $\gamma_{\text{XS}} = 0$ are closely related. In particular, it is straightforward to show that problem (2) is equivalent to applying the spectral decomposition to the covariance matrix of asset returns and problem (3) for the case with $\gamma_{\text{XS}} = 0$ is equivalent to applying the spectral decomposition to the second-moment matrix of asset returns ($R^\top R/T$). We rely on the formulation of PCA with uncentered returns to simplify notation and facilitate comparison with the existing literature (Lettau and Pelger, 2020b).

Figure 1: Cross-sectional shape target

This figure illustrates the cross-sectional shape target, $\Lambda^{\text{XS}} \in \mathbb{R}^{370\times 37}$, 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.



pricing characteristics are motivated by a univariate analysis that shows monotonic patterns in risk premia, aligned with certain firm characteristics (that is, single-sorted portfolios tend to produce a spread in expected returns). Thus, 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 decile 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 $N = 370$ assets, but also specifically long-short univariate risk premia. Second, the shape target nudges the loadings and factor weights to tilt towards this monotonic 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 enhances the interpretability of the statistical factors. While our analysis focuses on monotonic shapes, it is straightforward to include alternative shapes to capture for example convex 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 are

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 the innovations of $L = 7$ important macroeconomic time series. Given a panel of L macroeconomic time series Y^{macro}, we obtain the macroeconomic target as

$$\Lambda^{\text{XS}} = R^\top Y^{\text{macro}} \left((Y^{\text{macro}})^\top Y^{\text{macro}} \right)^{-1}. \tag{7}$$

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 to the second moment. Conventional PCA 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 considered. 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 one to estimate factors whose covariance signal is too weak to be detected by conventional PCA.

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:

$$\text{RP-PCA:} \quad \hat{F}_{\text{RP}}, \hat{\beta}_{\text{RP}} = \underset{F,\beta}{\text{argmin}} \quad \underbrace{\frac{1}{NT} \left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{variation of returns}}} + \gamma_{\text{RP}} \quad \underbrace{\frac{1}{N} \left\| \bar{R} - \bar{F}\beta^\top \right\|_F^2}_{\substack{\textbf{TS target}\text{: cross-sectional} \\ \text{pricing error}}}, \tag{8}$$

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

Our time-series target PCA (TS-target-PCA) approach 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 $(G^{\text{TS}})^\top R/T$, for a time-series target matrix $G^{\text{TS}} \in \mathbb{R}^{T \times J}$. The RP-PCA approach is the special case when the TS-target matrix is the T-dimensional vector of ones, $G^{\text{TS}} = \mathbf{1}_T$. The cross-sectional span of the loadings β is nudged to be aligned with the cross-sectional span of the target moments $(G^{\text{TS}})^\top R/T$. For this purpose, we introduce the *time-series*

projection matrix

$$P_G^{\text{TS}} = G^{\text{TS}} \left((G^{\text{TS}})^\top G^{\text{TS}} \right)^{-1} (G^{\text{TS}})^\top, \tag{9}$$

which projects T-dimensional vectors onto the column space of G^{TS}. The TS-target-PCA approach estimates latent factors using the following objective:

$$\text{TS-target-PCA:} \quad \hat{F}_{\text{TS}}, \hat{\beta}_{\text{TS}} = \underset{F,\beta}{\text{argmin}} \quad \underbrace{\frac{1}{NT} \left\| R - F\beta^\top \right\|_F^2}_{\substack{\text{unexplained} \\ \text{variation of returns}}} + \gamma_{\text{TS}} \underbrace{\frac{1}{NT} \left\| P_G^{\text{TS}} R - P_G^{\text{TS}} F\beta^\top \right\|_F^2}_{\substack{\textbf{TS target}\text{: moment error for} \\ P_G^{\text{TS}}\text{-projected returns}}}, \tag{10}$$

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

The factors that optimize the TS-target-PCA objective in (10) 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 optimize the TS-target-PCA objective can be estimated by applying PCA to the $N \times N$ matrix*

$$\frac{1}{NT} R^\top \left(I_T + \gamma_{TS} P_G^{TS} \right) R \tag{11}$$

Specifically, the estimated factor weights $\hat{\omega}_{TS}$ are the first K eigenvectors of matrix (11), and the factor returns are $\hat{F}_{TS} = R\hat{\omega}_{TS}$.

As the factor returns 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 R on $\hat{F}_{\text{TS}}$ 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 TS-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:

$$G^{\text{TS}} = \left(I_T - \tilde{F}^{\text{ref}} \left((\tilde{F}^{\text{ref}})^\top \tilde{F}^{\text{ref}} \right)^{-1} (F^{\text{ref}})^\top \right) \mathbf{1}_T, \tag{12}$$

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 $R^\top G^{\text{TS}}/T = \alpha^{\text{ref}}$, 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-alpha-PCA penalty nudges the latent factors to 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. Thus, 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 sets of constraints are non-redundant, and the combined estimator can outperform models that use only one of the targets. In addition, combining risk-premium penalty with the shape constraint enhances the interpretability of the factors.

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

$$\underset{F,\beta}{\text{argmin}} \underbrace{\frac{1}{NT}\left\|R - F\beta^\top\right\|_F^2}_{\substack{\text{unexplained}\\ \text{variation of returns}}} + \gamma_{\text{XS}} \underbrace{\frac{1}{LT}\left\|RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}}\right\|_F^2}_{\substack{\textbf{XS target}\text{: unexplained variation}\\ \text{of } P_\Lambda^{\text{XS}}\text{-projected returns}}}$$
$$+ \gamma_{\text{TS}} \underbrace{\frac{1}{NT}\left\|P_G^{\text{TS}}R - P_G^{\text{TS}}F\beta^\top\right\|_F^2}_{\substack{\textbf{TS target}\text{: moment error for}\\ P_G^{\text{TS}}\text{-projected returns}}} + \gamma_{\text{XS}}\cdot\gamma_{\text{TS}} \underbrace{\frac{1}{LT}\left\|P_G^{\text{TS}}RP_\Lambda^{\text{XS}} - P_G^{\text{TS}}F\beta^\top P_\Lambda^{\text{XS}}\right\|_F^2}_{\substack{\textbf{XS-TS target}\text{: } P_G^{\text{TS}}\text{-moment error}\\ \text{of } P_\Lambda^{\text{XS}}\text{-projected returns}}}. \tag{13}$$

The first term is the unexplained variation of returns, which is a conventional PCA-type objective. The second term is the cross-sectional penalty of XS-target-PCA and captures the unexplained variation in managed target portfolios returns 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 portfolio returns, 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-target-PCA objective in* (13) *are estimated by applying PCA to the* $N \times N$ *matrix*

$$\frac{1}{NT}\left(I_N + \tilde{\gamma}_{XS}P_\Lambda^{XS}\right)R^\top\left(I_T + \gamma_{TS}P_G^{TS}\right)R\left(I_N + \tilde{\gamma}_{XS}P_\Lambda^{XS}\right), \tag{14}$$

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* (14)*, then the estimated factor are* $\hat{F}_{XS\text{-}TS} = R\hat{\omega}^{TS\text{-}XS}$ *with the esti-*

mated factor weights

$$\hat{\omega}^{TS\text{-}XS} = \left(\frac{1}{\tilde{\gamma}_{XS}+1} I_N + \frac{\tilde{\gamma}_{XS}}{\tilde{\gamma}_{XS}+1} P_\Lambda\right) Q_K. \tag{15}$$

Given the estimated $\hat{F}_{\text{XS-TS}}$, we obtain consistent estimates of the loadings with an OLS regression of the returns on the factors. 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 estimators 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 estimators 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. 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 target matrix Λ^{XS} and the time-series target matrix G^{TS}. Then, we obtain the corresponding projection matrices P_Λ^{XS} and P_G^{TS}, 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 (14) 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

$$R_n = \alpha_n + \hat{F}_{\text{XS-TS}} \beta_n^\top + e_n, \tag{16}$$

   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 intuition why it is helpful to include targets in the estimation. The formal econometric theory and statistical argument for the XS-target-PCA is in Duan et al. (2023) 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.

The important effect of both the XS and TS targets 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. Roughly speaking, the eigenvalues of factors are boosted in the transformed matrix with the target information. This can have different beneficial effects. Weak factors, which cannot be detected from the covariance matrix alone, can now be estimated. Semi-weak factors, which affect a small but sufficiently large 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 moments 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. (2023). Assume that our panel is described by two factors $R = F\beta^\top + e$, 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

$$\Lambda^{\mathrm{XS}} = \begin{pmatrix} \underbrace{I_{N_1}}_{N_1\times N_1} \\ \underbrace{0}_{N_2\times N_1} \end{pmatrix}.$$

The target managed portfolios $R^{\mathrm{XS}} = R\Lambda^{\mathrm{XS}}$ implied by Λ^{XS} simply equal the first N_1 assets, which are only affected by factor one. The key variables in latent-factor estimation are the systematic eigenvalues from the factors given by

$$\frac{1}{NT} F\beta^\top \left(I_N + \gamma_{\mathrm{XS}} \frac{N}{N_1} \Lambda^{\mathrm{XS}} (\Lambda^{\mathrm{XS}})^\top \right) \beta F^\top$$

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

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 $\gamma_{\text{XS}} = r$ 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. (2023) 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 use 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 serves 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

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

should converge to the eigenvalues of the signal matrix

$$\Sigma_F + (1+\gamma_{\text{TS}})\,\mu_F \mu_F^\top. \tag{17}$$

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 matrix (17) can be pushed up. Thus, the factor can be detected or estimated more precisely.

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

$$\Sigma_F + \mu_F \mu_F^\top + \frac{\gamma_{\text{TS}}}{1+\text{SR}^2(F^{\text{ref}})} \alpha_{\text{ref}} \alpha_{\text{ref}}^\top,$$

where α_{ref} is the pricing error of the factors F relative to the reference factors F^{ref} and $\text{SR}^2(F^{\text{ref}})$ is the squared Sharpe ratio of the reference factors. Hence, we can push up the eigenvalues of weak factors with small variance, which 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

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

3.1 Data and evaluation methodology

We evaluate our method on a comprehensive set of characteristic-sorted portfolios. Specifically, 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 370 value-weighted decile portfolios obtained from 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 the Appendix lists the 37 characteristics and classifies them into eight categories.

Given factors $\hat{F}$ we can compute the maximum Sharpe ratio from the tangency portfolio of the mean-variance frontier that is spanned by $\hat{F}$ as

$$\hat{\omega}_{MV} = \Sigma_{\widehat{F}}^{-1} \mu_{\widehat{F}},$$

where $\mu_{\widehat{F}}$ and $\Sigma_{\widehat{F}}$ are the mean and covariance matrix of $\hat{F}$. The implied SDF is given by

$$M_t = 1 - \hat{\omega}_{MV}^{\top} \left(\hat{F}_t - \mu_{\widehat{F}} \right).$$

We compare models using three criteria: (i) the maximum Sharpe ratio that can be obtained by a linear combination of the factors (SR), (ii) the root-mean-square pricing error $\text{RMS}_\alpha = \sqrt{\hat{\alpha}^\top \hat{\alpha}/N}$, and (iii) the average idiosyncratic variance $\overline{\sigma}_e^2 = \frac{1}{N} \sum_{i=1}^{N} (\sigma_{\widehat{e}_i}^2 / \sigma_{R_i}^2)$. Unless otherwise mentioned, all reported results are out-of-sample using the following procedure: Factors and loadings are estimated on rolling windows of 20 years ($T = 240$). We use time-series regressions with intercept to obtain the loadings as described in Equation (16). With these estimated factor portfolio weights and loadings including information up to time t we obtain the out-of-sample factor returns and residuals at time $t + 1$. The mean and variance of the out-of-sample errors are used to calculate the average pricing error and the idiosyncratic variation, respectively. We use the optimal portfolio weights for the maximum Sharpe-ratio portfolio estimated in the rolling window period to create an out-of-sample optimal return giving us the maximum Sharpe-ratio portfolio out-of-sample. This is the same standard evaluation approach used among others in Lettau and Pelger (2020b). All quantities are based on a monthly frequency.

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. Thus, we expect the loadings on asset-pricing factors to be monotonic on these characteristics. To incorporate this economic restriction into the recovery of factor models, we consider a cross-sectional target matrix, $\Lambda^{\text{XS}} \in \mathbb{R}^{370\times 37}$, whose lth column contains the weights of a long-short portfolio of the lth characteristic, as illustrated in Figure 1.

Specifically, the weights are linear in the decile rank and centered at zero.[5] 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. Specifically, the shape target rewards positive weights for high characteristic deciles and negative weights for low deciles.

To measure how factor weights fit the pattern implied by the cross-sectional target, we define the cross-sectional metric of the ith factor as

$$\text{XS-Metric}(\hat{\omega}_i) = 100 \times \frac{\|P_{\Lambda}^{\text{XS}}\hat{\omega}_i\|_2}{\|\hat{\omega}_i\|_2}, \tag{18}$$

where $\hat{\omega}_i$ is the estimated factor weight vector and P_{Λ}^{XS} is the cross-sectional projection matrix defined in Equation (4) for the cross-sectional target matrix Λ^{XS}. This metric is motivated by the penalty in Equation (5). Thus, the closer the cross-sectional metric of a latent factor is to 100, the better its weights (and hence associated factor loadings) are spanned by the economically motivated cross-sectional target. A factor with a cross-sectional metric of 100 perfectly aligns with the target. Similarly, we define the cross-sectional metric of a factor model as the corresponding metric of the implied SDF weights.

3.2.1 Model performance

Figure 2 shows how cross-sectional shape PCA (XS-shape-PCA) improves the performance over conventional PCA. We report the out-of-sample maximum Sharpe ratio, pricing errors, and unexplained variation for different numbers of factors and different values of the penalty parameter $\gamma_{\text{XS}}^{\text{shape}}$.

Panels A and B in Figure 2 show that the cross-sectional shape target helps to substantially increase the maximum Sharpe ratios of models with five or more factors. The case of $\gamma_{\text{XS}}^{\text{shape}} = 0$ corresponds to conventional PCA. The Sharpe ratio of a five-factor model with $\gamma_{\text{XS}}^{\text{shape}} = 1$ increases by an impressive 57.4% (from 0.18 to 0.28) relative to the PCA case. Interestingly, the Sharpe ratio does not vary substantially with the penalty parameter for models with one to four factors. This suggests that the fifth XS-shape-PCA factor is a weak factor that is missed by conventional PCA

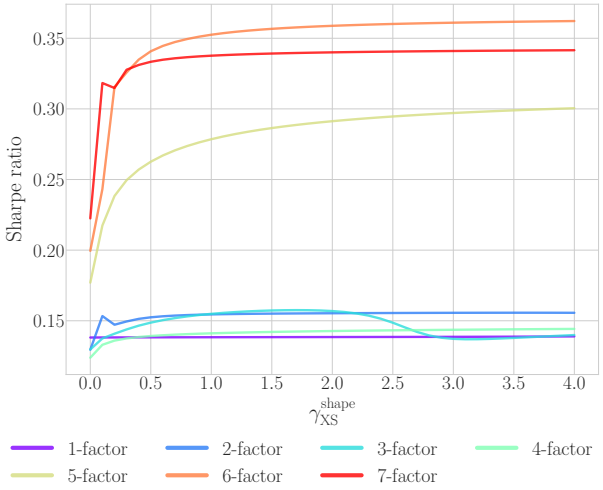
[5] Mathematically, the element on the nth row and lth column of Λ^{XS} is

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

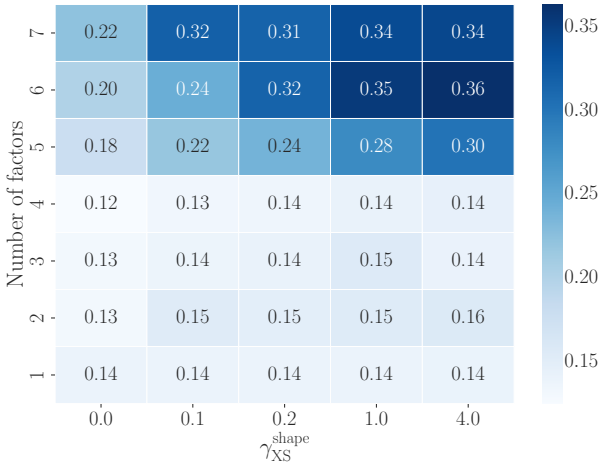
Figure 2: Out-of-sample performance of cross-sectional shape PCA models

This figure shows the out-of-sample performance of the cross-sectional shape PCA (XS-shape-PCA) models for different numbers of factors and different values of the penalty parameter $\gamma_{\text{XS}}^{\text{shape}}$. Panel A and Panel B report the maximum Sharpe ratio (SR) of the tangency portfolio for up to seven factors. Panel C displays the root-mean-squared pricing error RMS_α of the top-minus-bottom portfolios of the 37 characteristics. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.

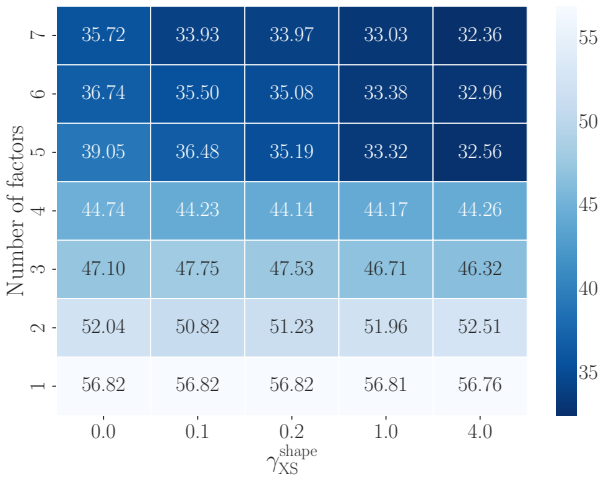
(A) Sharpe ratio (SR, line plot)



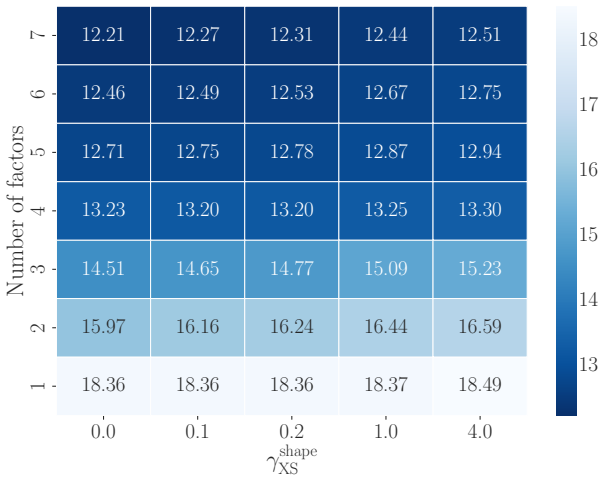
(B) Sharpe ratio (SR, heatmap)



(C) RMS_α: top-minus-bottom portfolios



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



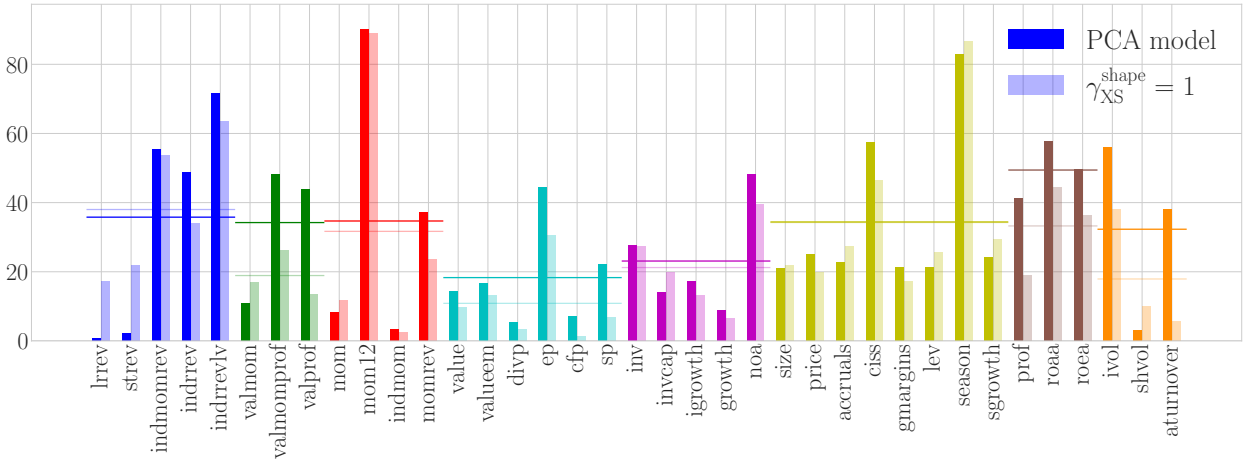
but is important for cross-sectional asset pricing. Increasing the penalty parameter on the shape target enhances the signal of this weak factor, which allows XS-shape-PCA to capture it.

Panel C demonstrates that the shape target reduces the out-of-sample pricing errors of target assets that incorporate monotonicity information. We report the root-mean-squared pricing error of 37 top-minus-bottom portfolios. Each of these portfolios is a long-short portfolio in the extreme deciles of a characteristic. Thus, these portfolios closely mimic the shape target portfolios, which are univariate long-short portfolios whose weights are linear in the ranks. The results are consistent with the Sharpe ratio findings. In particular, the pricing error of a five-factor model

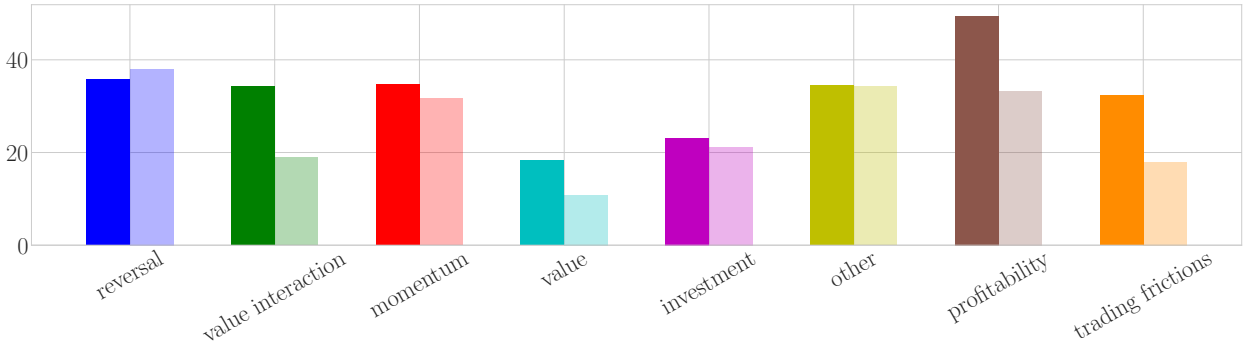
Figure 3: Pricing errors for top-minus-bottom portfolios for XS-shape-PCA

This figures shows the out-of-sample pricing errors for the 37 top-minus-bottom portfolios of the five-factor PCA model and XS-shape-PCA model with $\gamma_{\text{XS}}^{\text{shape}} = 1$. Panel A shows the pricing errors for all 37 portfolios, whereas Panel B graphs average pricing error for each anomaly category. The light and dark bars correspond to XS-shape-PCA and PCA, respectively. Pricing errors are reported in basis points.

(A) Pricing error of each portfolio



(B) Average pricing error of portfolios in each category



with $\gamma_{\text{XS}}^{\text{shape}} = 1$ falls to 33.32 basis points (a 14.7% reduction) relative to the conventional PCA benchmark with 39.05. As expected, the average pricing error of a model with less than five factors does not vary substantially with the penalty parameter.

Panel D demonstrates that the explained variation is not much affected by the cross-sectional shape penalty. As expected, PCA yields the lowest average unexplained variance, as it is designed to explain the time-series variation of test assets. However, the average unexplained variance essentially does not change with the penalty parameter. Hence, XS-shape-PCA factors result in better asset pricing results while explaining a similar amount of time-series variation as PCA factors. Panel D also confirms that the fifth factor is a weak factor. In more detail, the fifth PCA factor explains only $13.23\% - 12.71\% = 0.52\%$ of the time-series variation, whereas the fifth XS-shape-PCA factor with $\gamma_{\text{XS}}^{\text{shape}} = 1$ explains $13.25\% - 12.87\% = 0.38\%$. However, the fifth XS-shape-PCA factor is substantially more relevant for asset pricing.

Based on these results, we set $\gamma_{\text{XS}}^{\text{shape}} = 1$ as the benchmark for XS-shape-PCA. Figure 2 shows that increasing $\gamma_{\text{XS}}^{\text{shape}}$ beyond one does not substantially change the performance of the models. As the five-factor model captures a substantial amount of the pricing information in terms of pricing errors and Sharpe ratio, we focus our discussion in the main text on the five-factor XS-shape-PCA model with $\gamma_{\text{XS}}^{\text{shape}} = 1$. The appendix and internet appendix report the complete analysis for other parameters.

We now compare in detail how XS-shape-PCA and conventional PCA models price the top-minus-bottom portfolios of characteristics. Figure 3 shows the pricing errors for the 37 top-minus-bottom portfolios of the five-factor XS-shape-PCA with $\gamma_{\text{XS}}^{\text{shape}} = 1$ and conventional PCA. Panel A demonstrates that XS-shape-PCA delivers a lower pricing error than PCA for 26 of the 37 top-minus-bottom portfolios. Importantly, XS-shape-PCA performs particularly well in pricing the portfolios that the conventional PCA finds hard to price. When aggregating the top-minus-bottom portfolios into different characteristic categories in Panel B, we find that XS-shape-PCA delivers a lower average pricing error than conventional PCA for every category except for reversal, where both models are en par. In particular, the shape target leads to a reduction of more than 40% in the average pricing error for the value-interaction, profitability, and trading-frictions categories.

Overall, we conclude that the cross-sectional shape target generally improves the pricing performance over the conventional PCA model by capturing weak factors that are relevant for asset pricing. In particular, the Sharpe ratio and pricing errors for the top-minus-bottom portfolios of characteristics improve substantially with the penalty on the shape target, while the explained time-series variation essentially stays the same.

3.2.2 Factor structure

Next, we study the compositions of the XS-shape-PCA factors and compare them with the PCA factors. Figure 4 shows the portfolio weights on all decile portfolios for the fifth factors of XS-shape-PCA with $\gamma_{\text{XS}}^{\text{shape}} = 1$ and PCA. As shown in the previous section, the weak fifth factor is one of the main drivers for the better pricing performance of XS-shape-PCA. We observe a clear monotonic pattern for XS-shape-PCA, while PCA has a more diffuse pattern. The monotonic shape is particularly pronounced for characteristics in the reversal, value-interaction, profitability, and trading-frictions categories. We conclude that a monotonic shape constraint is a useful economic prior that helps to discipline latent factor recovery.

The more interpretable monotonic shape in the factor portfolio weights holds for all factors. Figure C.2 reports the corresponding results for the remaining factors. While PCA factors generally have a diffuse pattern, XS-shape-PCA yields a clear monotonic shape structure for all the factors. The only exception is the first factor, which, as expected, is a typical long-only market factor.

A shape target not only affects the monotonic pattern of the portfolio weights, but also which characteristics receive higher weights. Figure 5 depicts the weights on the top and bottom characteristic deciles of each of the first six XS-shape-PCA factors. For comparison, Figure C.1 shows the weights for PCA factors. The characteristics are grouped into eight categories with

Figure 4: Portfolio weights of fifth XS-shape-PCA factor

This figure shows the weights on all 370 decile portfolios for the fifth factors of XS-shape-PCA with $\gamma_{\text{XS}}^{\text{shape}} = 1$ and PCA. Positive (negative) weights are shown in green (red). The decile portfolios are grouped by characteristics. Figure C.2 shows the weights for other factors one to six. The factor weights are estimated on the full sample.

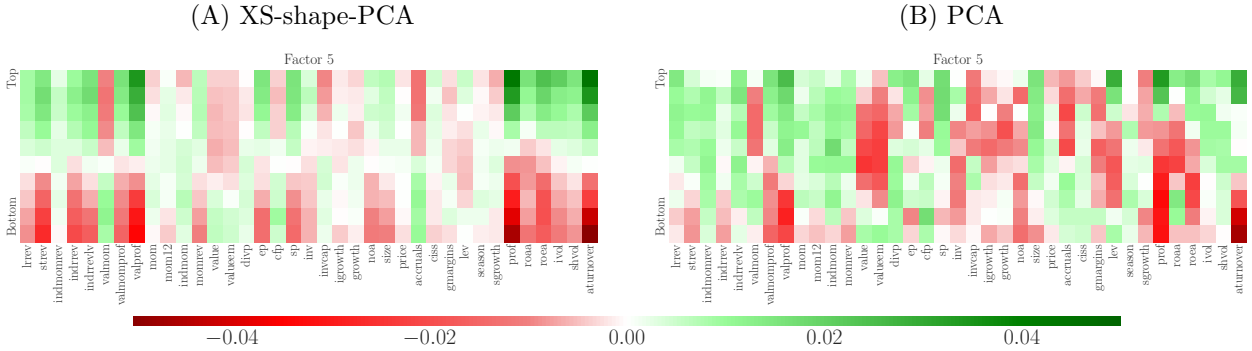
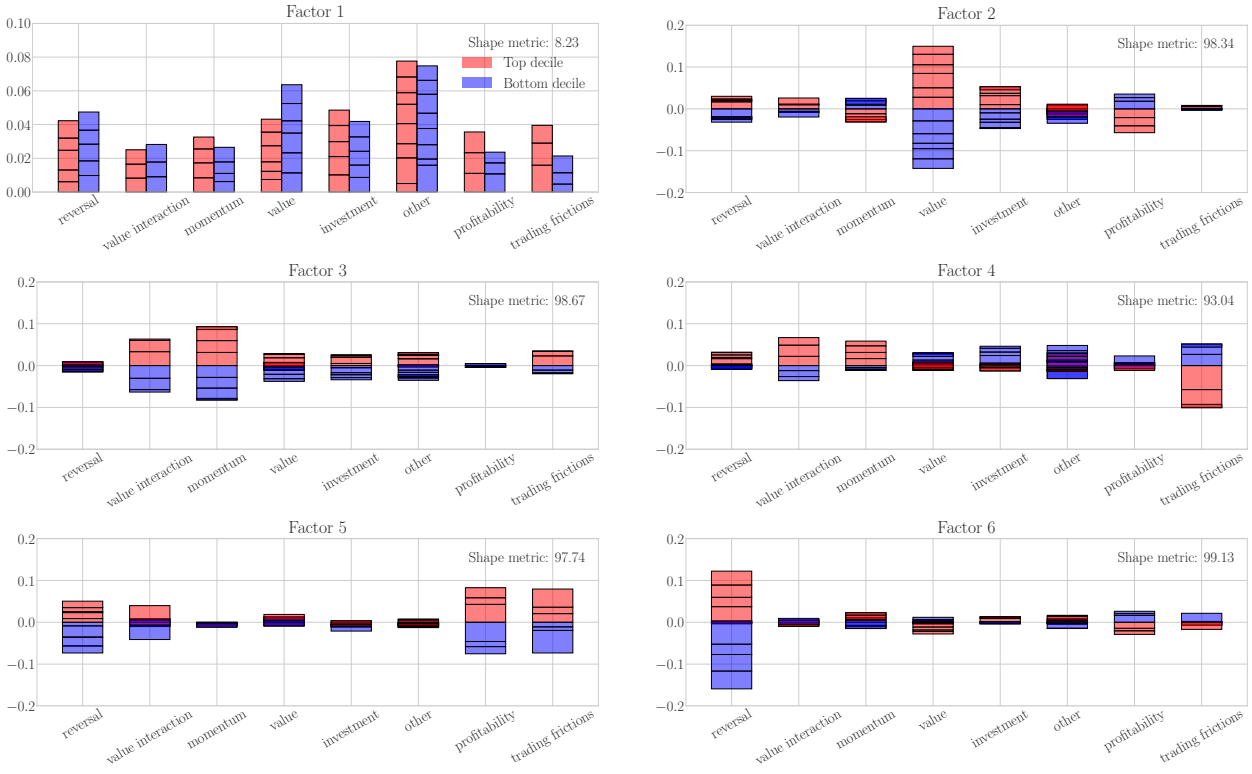


Figure 5: XS-shape-PCA factor weights on top- and bottom-decile portfolios

This figure shows the portfolio weights on the top and bottom characteristic deciles of each of the first six XS-shape-PCA factors with $\gamma_{\text{XS}}^{\text{shape}} = 1$. The characteristics are grouped into eight categories. The vertical axis shows the weights on the top (bottom) deciles in red (blue). Each bar shows the total weight on the top (bottom) deciles aggregated across all characteristics in a category, with black lines separating the contribution of each characteristic in that category. All factors are normalized to have positive average returns. The figure also reports the cross-sectional shape metric of each factor. The factor weights are estimated on the full sample.



weights on the top deciles in red and bottom deciles in blue. Each bar aggregates the weights

across all characteristics in a category. We also report the cross-sectional shape metric of the factors in each panel.

Figure 5 shows that the cross-sectional shape target nudges the factor weights of *all* six XS-shape-PCA factors (with the exception of the first) to be monotonic on the characteristics. Specifically, the cross-sectional shape metric averaged across the second to sixth factors is 97.38 for XS-shape-PCA, but only 73.16 for conventional PCA. This demonstrates that XS-shape-PCA consistently incorporates the economic prior underlying the shape constraints into the construction of the factors because the cross-sectional shape metric measures how well the factor composition satisfies the long-short pattern.

The factors have a clear economic interpretation. The first XS-shape-PCA factor assigns similar positive weights to the top and bottom deciles, resembling a market factor. Indeed, the correlation of the first XS-shape-PCA factor with the market is above 0.99. The second XS-shape-PCA factor is closely related to value. The third factor assigns large long-short weights to portfolios in the momentum and value-interaction categories. The fourth factor puts most of its weight on trading frictions, but generally has a less clear pattern. The fifth factor has large long-short weights on the top and bottom deciles in the reversal, value-interaction, profitability, and trading-frictions categories. The sixth factor is closely related to reversal because it is almost exclusively composed of long-short reversal portfolios. This structure is distinctively different from PCA factors, which largely neglect the reversal and trading friction categories, and hence result in weaker pricing performance.

3.3 Time-series alpha target

We now consider a *time-series alpha target* that nudges latent factors to explain the risk premia that a candidate asset-pricing model fails to explain. We consider the five-factor Fama-French model (Fama and French, 2015, FF5) as the candidate model. Thus, we define the time-series target matrix G^{TS} as in Equation (12) with F^{ref} as the FF5 factor returns. As a motivating example, we first illustrate the effect of an alpha penalty with a small cross section containing the 25 double-sorted portfolios formed on size and accruals from Ken French's website. We then present our main results for TS-alpha-PCA on the larger cross section of 370 decile portfolios.

3.3.1 Motivating example: Size and accruals portfolios

As a motivating example, we start with the 25 double-sorted portfolios formed on size and accruals. Figure 6 displays the time-series alphas of the 25 double-sorted portfolios with respect to FF5. Figure 6 demonstrates that the FF5 model struggles to price high-accruals portfolios, with four out of the five high-accruals portfolios delivering negative alpha exceeding -15.70 basis points per month. In addition, the ME5-AC1 and ME1-AC2 portfolios have positive FF5 alphas above 23.63 basis points.

Figure 6: Alphas of 25 size and accruals portfolios for FF5

This figure shows the in-sample alphas with respect to the FF5 model (Fama and French, 2015) of the 25 size and accruals portfolios. The horizontal and vertical axes rank the portfolios from low accruals (AC1) to high accruals (AC5) and from small (ME1) to large (ME5), respectively. Positive (negative) alphas are shown in green (red) and all alphas are reported in basis points per month.



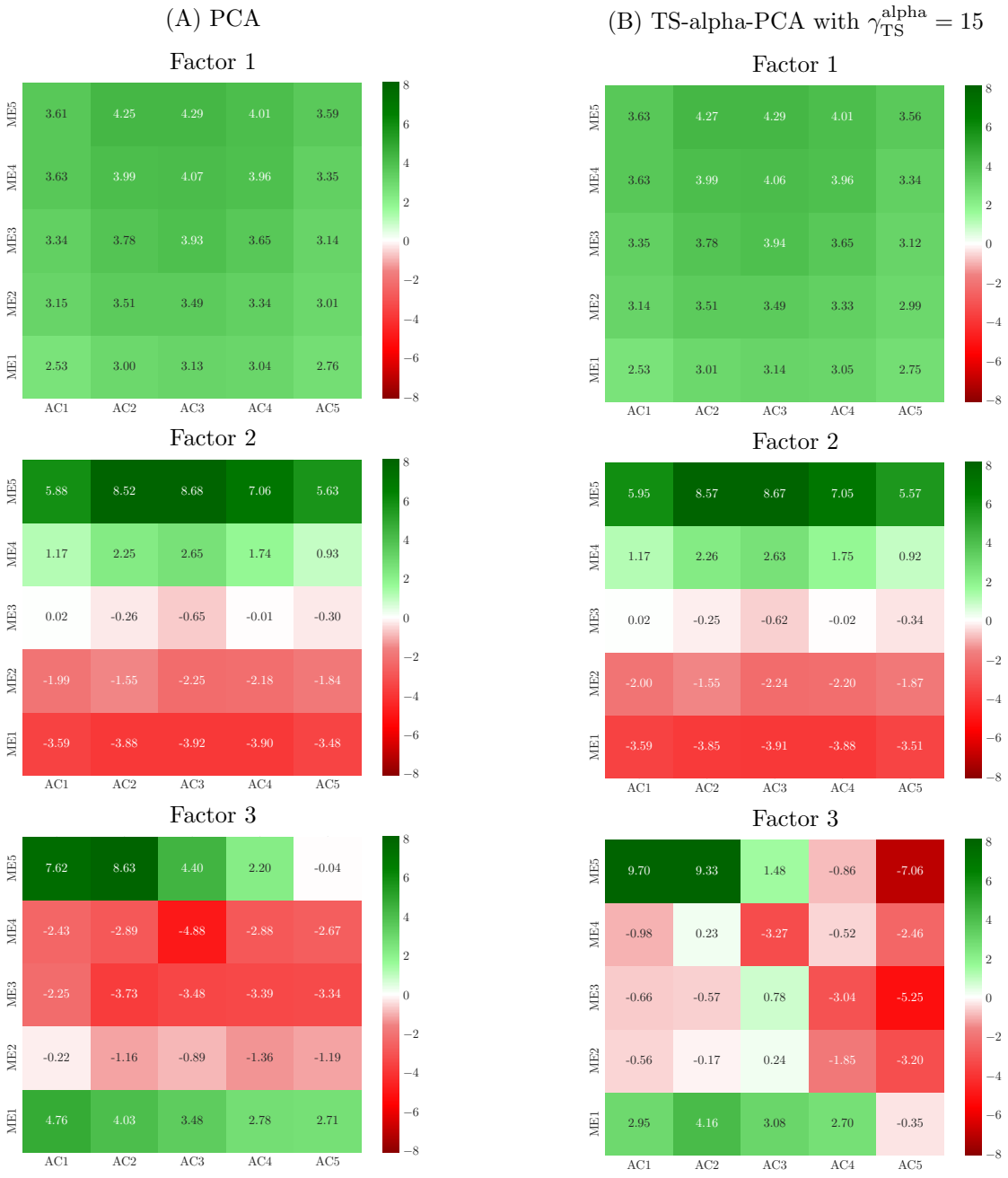
To illustrate how the time-series alpha target changes the compositions of the PCA factors, Figure 7 shows the factor weights on the 25 portfolios of the PCA factors and TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. The compositions of the first two PCA and TS-alpha-PCA factors are almost identical: the first factor assigns positive weights to all 25 portfolios and hence mimics a level factor, and the second factor mimics a size factor that assigns positive weights to large stocks and negative weights to small stocks.

In contrast, the structure of the third factor is quite different for PCA and TS-alpha-PCA. Importantly, the structure of the third TS-alpha-PCA factor reflects the pattern of the alphas on the test assets, illustrated in Figure 6. In particular, while the third PCA factor is a "curvature" factor in the size dimension, the third TS-alpha-PCA is a "FF5-alpha factor." The third TS-alpha-PCA factor assigns a large negative weight of -7.06 to the ME5-AC5 portfolio, which has a large negative alpha of -18.58, while the third PCA factor assigns a weight of only -0.04 to this portfolio. Similarly, the third TS-alpha-PCA factor assigns more negative weights (compared to the third PCA factor) to other portfolios with large negative FF5 alphas (ME1-AC5, ME2-AC5, and ME3-AC5) and more positive weights to portfolios with large positive alphas (ME5-AC1, ME1-AC2, and ME3-AC3).

The third TS-alpha-PCA factor better captures the risk premium information contained in the alphas of the size and accruals portfolios. The out-of-sample Sharpe ratio of the three-factor TS-alpha-PCA model is 0.20, which is substantially higher than that of its PCA counterpart, 0.14. As the first two factors of both models are almost identical, the increase in Sharpe ratio is driven by the third TS-alpha-PCA factor. Thus, TS-alpha-PCA better identifies the risk premia in the

Figure 7: TS-alpha-PCA factor weights on 25 size and accruals portfolios

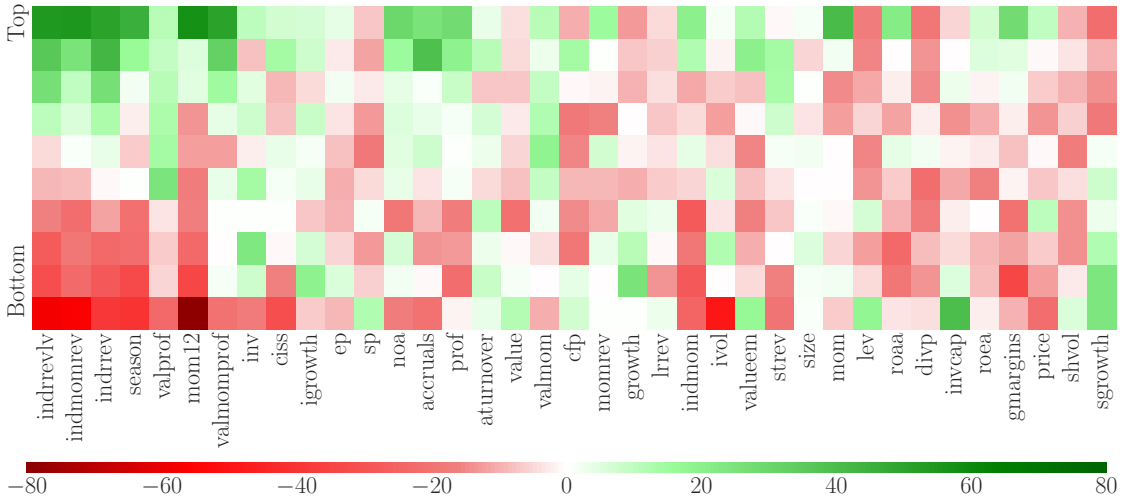
This figure shows the in-sample weights on the 25 size and accruals portfolios of PCA factors and TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. Each panel includes three subfigures for the first three factors. In each heatmap, the horizontal and vertical axes rank the portfolios from low accruals (AC1) to high accruals (AC5) and from small (ME1) to large (ME5). Positive (negative) weights are shown in green (red).



cross section by increasing the signals of factors that align with the alphas of the size and accruals portfolios, which capture the return spread not explained by FF5.

Figure 8: Alphas of all 370 decile portfolios for FF5

This figure shows the in-sample alphas of the 370 decile portfolios with respect to FF5 (Fama and French, 2015) in basis points per month. The horizontal axis ranks the characteristics in descending order of Sharpe ratios of top-minus-bottom portfolios. The vertical axis depicts the bottom- to top-decile portfolios. Positive (negative) alphas are shown in green (red).



3.3.2 Main results: The 370 decile portfolios

Next, we study the performance of time-series alpha PCA on our main cross section of 370 decile portfolios. Figure 8 displays the in-sample alphas of all 370 decile portfolios with respect to the FF5 model. The horizontal axis ranks the characteristics in descending order of top-minus-bottom portfolio Sharpe ratios. The figure demonstrates that the FF5 alphas of the decile portfolios, especially those of characteristics with high top-minus-bottom portfolio Sharpe ratios, have a strong monotonic pattern. The characteristics that the FF5 model struggles to price include reversal and momentum characteristics, among others. Overall, Figure 8 shows that FF5 cannot fully explain the cross section of decile portfolios. Thus, there exists a risk premium component for this cross section that FF5 fails to capture, and we expect time-series alpha PCA to help to identify factors that capture such risk premium information missed by FF5.

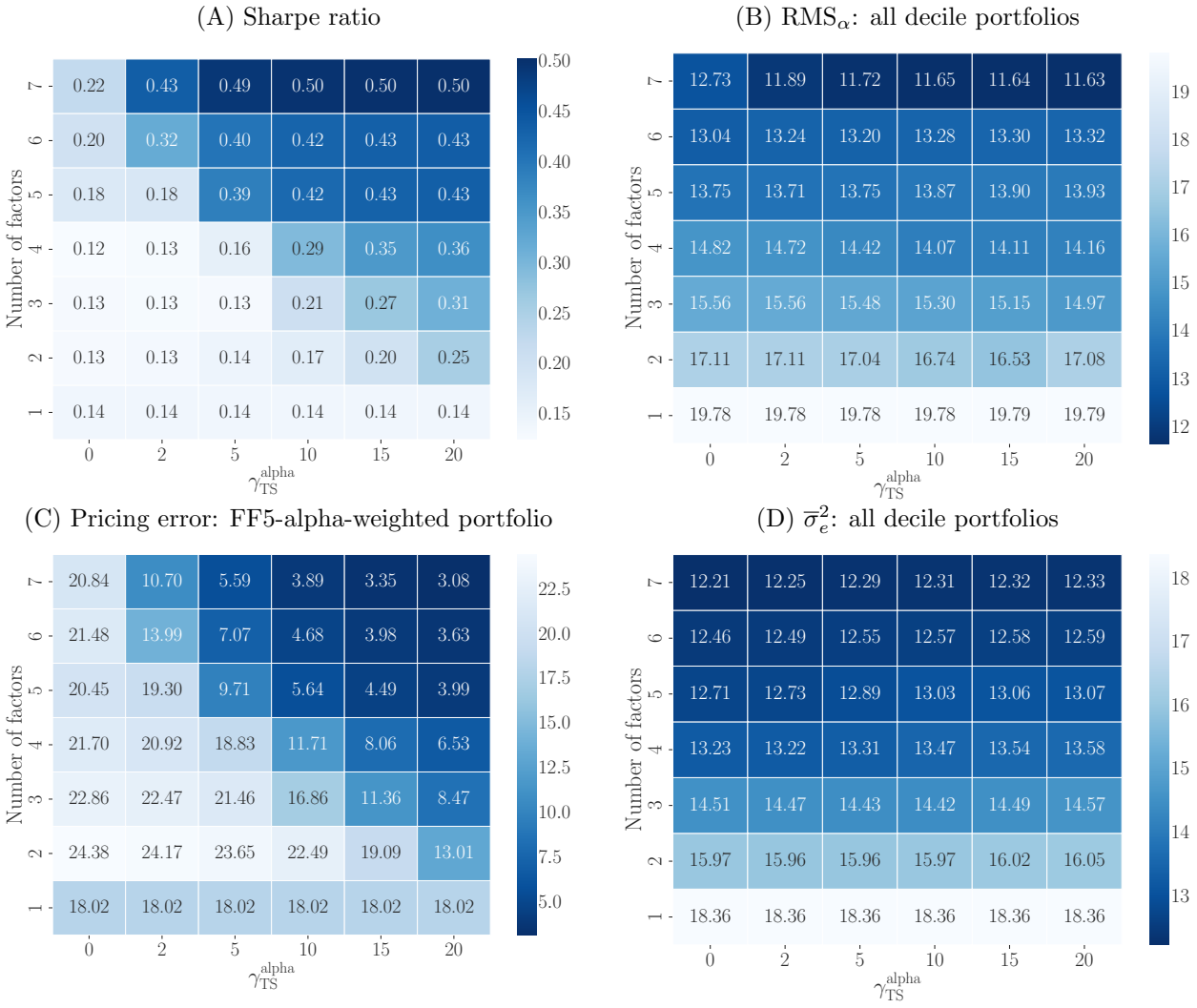
Model performance

Figure 9 shows how time-series alpha PCA (TS-alpha-PCA) improves the out-of-sample performance over conventional PCA. We report the out-of-sample maximum Sharpe ratio, pricing errors for different test assets, and the unexplained variation for different numbers of factors and values of the penalty parameter $\gamma_{\text{TS}}^{\text{alpha}}$.

Panel A of Figure 9 shows that the time-series alpha target substantially increases the maximum Sharpe ratio of the implied tangency portfolio. The Sharpe ratio of the five-factor model more than doubles from 0.18 to 0.43 when increasing the penalty parameter from zero (conventional PCA) to 15. Even for the three-factor model we observe an increase from 0.13 to

Figure 9: Out-of-sample performance of time-series alpha PCA models

This figure shows the out-of-sample performance of time-series alpha PCA (TS-alpha-PCA) for different numbers of factors and different values of the penalty parameter $\gamma_{\text{TS}}^{\text{alpha}}$. Panel A reports the maximum Sharpe ratio (SR) of the tangency portfolio for up to seven factors. Panel B shows the root-mean-squared pricing error RMS_α for the 370 decile portfolios. Panel C displays the pricing error for the *FF5-alpha-weighted portfolio*, which is a portfolio where the decile weights are proportional to the decile's alpha with respect to FF5. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



0.27 for $\gamma_{\text{TS}}^{\text{alpha}} = 15$. This is evidence that the time-series alpha target helps to capture risk premium information. Increasing the penalty on the time-series alpha target enhances the signals of factors that carry risk premia not captured by the FF5 model so that they are identified as lower-order TS-alpha-PCA factors, and thus, the TS-alpha-PCA consistently delivers higher Sharpe ratios than conventional PCA.

Given that the time-series alpha target helps to increase the Sharpe ratio, one may expect that it will also help to reduce the average pricing error across all decile portfolios. However, Panel B shows that the time-series alpha target does not substantially affect the root-mean-squared

pricing error across the decile portfolios. This is because the time-series alpha target focuses on capturing risk premia unexplained by FF5, but not on minimizing the average pricing error of all decile portfolios. In other words, the objective of the time-series alpha PCA is *not* to explain the expected returns of the decile portfolios, but rather to explain the *alphas* of the decile portfolios with respect to the FF5 model. As this is a different objective, there is no reason why the average pricing error has to decrease for higher time-series alpha targets. It is good to keep in mind that an average pricing error is a counting metric, while the Sharpe ratio is a spanning metric. The time-series alpha target focuses on pricing a subset of the large number of decile portfolios. As this subset has high risk premia, the implied factors can increase the Sharpe ratio. Next, we will look at a portfolio that focuses on this subset of test assets with large alphas.

Panel C shows that the time-series alpha target helps to substantially decrease the pricing error for the FF5-alpha-weighted portfolio, which is designed to capture risk premia unexplained by the FF5 model. The FF5-alpha-weighted portfolio is an investable portfolio using past FF5 time-series alphas as portfolio weights. Specifically, it is a long-short portfolio, where the positive and negative weights are normalized to one and proportional to the positive, respectively negative, FF5-alphas. The FF5-alpha-weighted portfolio has an out-of-sample average return of 19.5 basis points. A five-factor PCA model yields a large pricing error of 20.45 basis points, failing to explain any pricing information for this portfolio. In contrast, TS-alpha-PCA with $\gamma_{\text{TS}}^{\text{alpha}} = 15$ and five factors achieves an astonishing pricing error of 5.64. This demonstrates that TS-alpha-PCA achieves its goal to capture risk premia unexplained by FF5 and explain the mean return of the FF5-alpha-weighted portfolio.

Panel D confirms, as before, that the unexplained variation is not much affected by a target penalty. As expected, a conventional PCA model generally delivers a lower average unexplained variance, but TS-alpha-PCA achieves a comparable result. This is in line with our interpretation that TS-alpha-PCA helps to detect weak factors with pricing information, while also capturing the strong factors that explain the time-series variation.

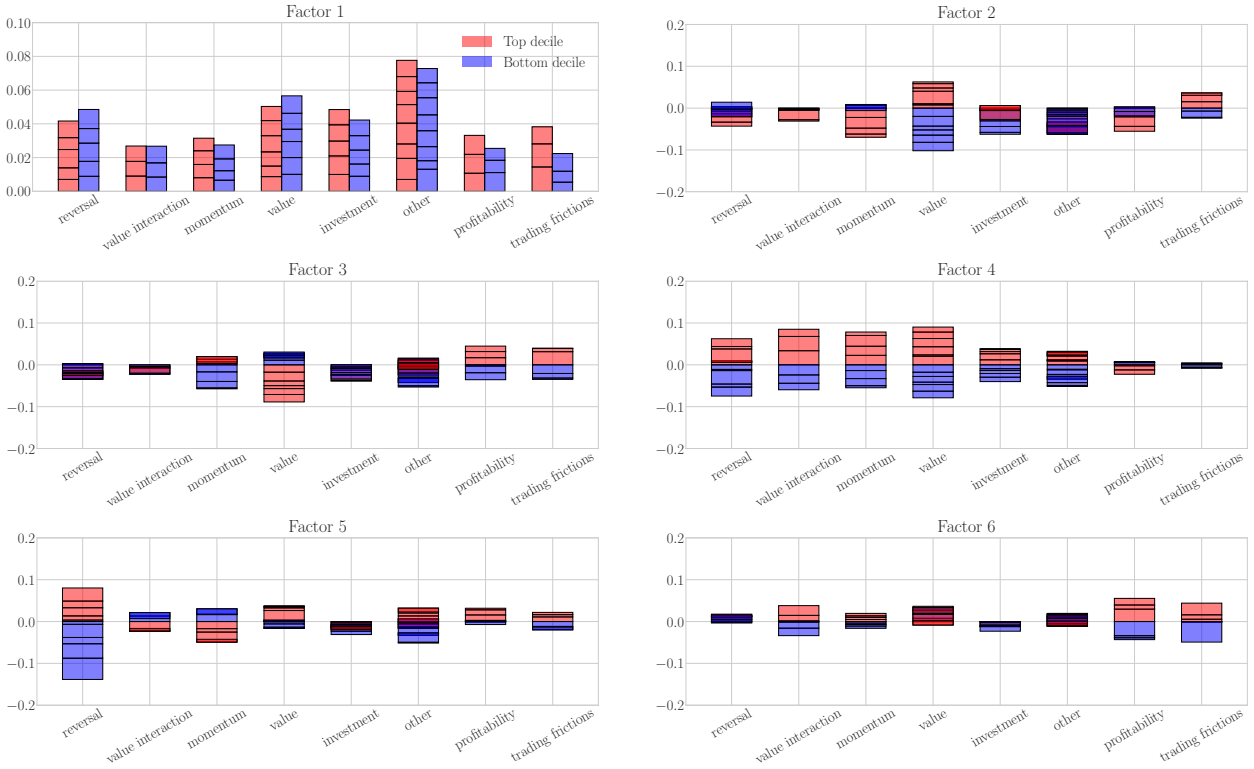
Based on these results, we set $\gamma_{\text{TS}}^{\text{alpha}} = 15$ as the benchmark for TS-alpha-PCA. Figure 9 shows that increasing $\gamma_{\text{TS}}^{\text{alpha}}$ beyond 15 does not substantially change the performance of the models. As the five-factor model captures a substantial amount of the pricing information in terms of pricing errors and Sharpe ratio, we focus our discussion in the main text on the five-factor TS-alpha-PCA model with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. The appendix and internet appendix report the complete analysis for other parameters.

Factor structure

We now study the compositions of the TS-alpha-PCA factors and compare them with the PCA factors. Figure 10 shows the weights on the top and bottom characteristic deciles of the first six TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. For comparison, Figure C.3 shows the weights for the PCA factors. The characteristics are grouped into eight categories with weights on the top deciles in red and bottom deciles in blue. Each bar aggregates the weights across all characteristics in a category.

Figure 10: TS-alpha-PCA factor weights on top- and bottom-decile portfolios

This figure shows the portfolio weights on the top and bottom characteristic deciles of each of the first six TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$. The characteristics are grouped into eight categories. The vertical axis shows the weights on the top (bottom) deciles in red (blue). Each bar shows the total weight on the top (bottom) deciles aggregated across all characteristics in a category, with black lines separating the contribution of each characteristic in that category. All factors are normalized to have positive average returns. The factor weights are estimated on the full sample.



Comparing the compositions of the TS-alpha-PCA factors with those of conventional PCA factors, we find that the fourth and fifth TS-alpha-PCA factors have a long-short pattern on characteristics in the reversal category that is missing in conventional PCA factors. This is not surprising as portfolios in the reversal category, such as *indrrevlv*, *indmomrev*, and *indrrev*, deliver high Sharpe ratios and have large alphas as demonstrated in Figure 8. Hence, the weaker fourth and fifth factors are substantially different for TS-alpha-PCA. TS-alpha-PCA detects the reversal risk premia signal, which substantially increases the Sharpe ratio of the implied SDF. Conventional PCA puts a larger weight on the value category, which is well explained by FF5 and hence receives a lower weight for TS-alpha-PCA.

Time-series mean target

We consider the important case of a time-series mean target. This can be considered as a special case of TS-alpha-PCA with a zero-factor candidate model. TS-mean-PCA targets to explain the mean returns of the full cross-section and not the alphas relative to a candidate factor model. This

model has been introduced and studied in Lettau and Pelger (2020b). To stay consistent with our notation, we use the label TS-mean-PCA, but this should be understood as a synonym for RP-PCA. The appendix provides the complete analysis for TS-mean-PCA, where we show that we can replicate the main findings of Lettau and Pelger (2020b). Here, we provide a brief summary of the results.

Figure D.8 shows how TS-mean-PCA improves the out-of-sample performance over conventional PCA. We report the out-of-sample maximum Sharpe ratio, pricing errors for different test assets, and the unexplained variation for different numbers of factors and values of the penalty parameter $\gamma_{\text{TS}}^{\text{mean}}$. The mean target substantially increases the maximum Sharpe ratio of the implied tangency portfolio. The Sharpe ratio of the five-factor TS-mean-PCA with $\gamma_{\text{TS}}^{\text{mean}} = 15$ is 0.46, which is almost three times larger than the conventional five-factor PCA model with 0.18. The objective to minimize the average pricing error is reflected in the RMS_α for the full cross-section. The five-factor TS-mean-PCA decreases the RMS_α from 13.75 to 11.84 basis points relative to the PCA benchmark. As expected the explained variation is largely unaffected by the mean penalty.

Figure C.4 shows the compositions of the TS-mean-PCA factors. The weak fifth factor is a major driver of the strong performance of TS-mean-PCA. It puts a large weight on the reversal category, which is missed by conventional PCA. Figure C.5 compares the factor composition with TS-alpha-PCA for FF5. Factor five is similar for TS-alpha-PCA and TS-mean-PCA, that is, both methods detect this weak factor with a high Sharpe ratio. However, the third and fourth factors differ for these methods. TS-mean-PCA puts more weight on the value and investment categories. Not surprisingly, TS-alpha-PCA puts less weight on these categories as those are relatively well explained by FF5.

We set $\gamma_{\text{TS}}^{\text{mean}} = 15$ as the benchmark for TS-mean-PCA, as increasing $\gamma_{\text{TS}}^{\text{mean}}$ beyond 15 does not have any substantial effect. Similar to the other methods, we focus our discussion in the main text on the five-factor TS-mean-PCA model with $\gamma_{\text{TS}}^{\text{mean}} = 15$, as it captures a substantial amount of the pricing information in terms of pricing errors and Sharpe ratio.

Out-of-sample performance summary

We compare the performance of different time-series targets and show that both, mean and alpha targets, substantially improve models, but focus on improving the pricing of different test assets. The targets are complementary and are not substitutes.

Table 1 summarizes the out-of-sample performance of TS-alpha-PCA, TS-mean-PCA, and conventional PCA. We also include the three-factor Fama and French (1993) and five-factor Fama and French (2015) models as benchmarks. We consider three and five latent factors, as based on the analysis in the previous sections, these are parsimonious representations that capture a substantial amount of pricing and variance information. This also provides a fair comparison in terms of the number of factors with the Fama-French three- and five-factor models. The benchmark penalty parameter on the time-series target is $\gamma_{\text{TS}} = 15$ for both, TS-mean-PCA and TS-alpha-PCA. We report the maximum Sharpe ratio of the tangency portfolio, RMS_α and unexplained variance for all

Table 1: Out-of-sample performance summary for time-series target PCA

This table summarizes the out-of-sample 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 $\gamma_{\text{TS}} = 15$ for TS-mean-PCA and TS-alpha-PCA. For each model we report the maximum 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. We calculate the time-series alpha metric of every model in each rolling window and report its time-series average. 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-weighted portfolio	$\overline{\sigma}_e^2$ for all decile portfolios	TS-alpha metric
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**

decile portfolios, and the pricing error for the FF5-alpha-weighted portfolio. In addition, we report a time-series alpha metric that measures how well the factors are aligned with the time-series alpha target.

Specifically, the time-series alpha metric is motivated by the objective function of TS-alpha-PCA in Equation (10). We define the time-series metric of K latent factors as

$$\text{TS-Metric} = 100 \times \frac{\|P_G^{\text{TS}} R P_{\hat{\omega}}\|_F}{\|P_G^{\text{TS}} R\|_F}, \tag{19}$$

where P_G^{TS} is the projection on the time-series target and $P_{\hat{\omega}}$ is the projection on the factor portfolio weights $\hat{\omega}$.[6] The time-series metric measures how well the factors align with the time-series target. In the case of the time-series alpha metric, the denominator of (19) simply measures the FF5 alphas of all decile portfolios, while the numerator measures the component of the FF5 alphas captured by the factors with portfolio weights $\hat{\omega}$. A time-series metric of 100 implies that the factors explain the full cross-section of FF5 alphas.

Table 1 demonstrates that both time-series targets (mean and alpha) substantially improve the Sharpe ratio of the implied tangency portfolio. For both, the three- and five-factor models, TS-mean-PCA and TS-alpha-PCA achieve considerably higher Sharpe ratios than PCA or the corresponding Fama-French benchmark. This indicates that both time-series targets capture more

[6] $P_{\hat{\omega}} = \hat{\omega}(\hat{\omega}^\top \hat{\omega})^{-1} \hat{\omega}^\top$ is an $N \times N$ matrix that projects the cross-section of returns onto the column space of the factor portfolio weights $\hat{\omega}$. Intuitively, it holds $RP_{\hat{\omega}} \approx F\beta^\top$.

risk premium information than the benchmarks. However, the two TS-target-PCA methods focus on explaining different test assets.

Time-series *alpha* PCA performs best at explaining the expected return of the FF5-alpha-weighted portfolio. This is consistent with the high time-series alpha metric, which measures how well the full cross-section of alphas is explained. This is because the alpha target is designed to explain the FF5 alphas of the decile portfolios and thus captures risk premia missed by FF5. As expected, the Fama-French factors perform badly at explaining this test asset. Surprisingly, conventional PCA is not much better than the Fama-French factors.

Time-series *mean* PCA models are best at explaining the average mean return of the full cross-section, which is reflected in the lowest RMS_α for all deciles. This is exactly the objective function of TS-mean-PCA. However, the pricing error on a specific subset of test assets, namely the decile portfolios with high FF5 alphas, is considerably larger than for TS-alpha-PCA. Hence, in spite of very similar Sharpe ratios for TS-mean-PCA and TS-alpha-PCA, their factors capture complementary risk premium information, which explains the returns of different test assets.

In summary, time-series target PCA models are superior to conventional PCA for asset pricing and outperform popular risk factors. Both of our time-series targets have economic value. Different targets in time-series target PCA optimize models for different objectives and discover weak factors that explain the risk premia of different test assets.

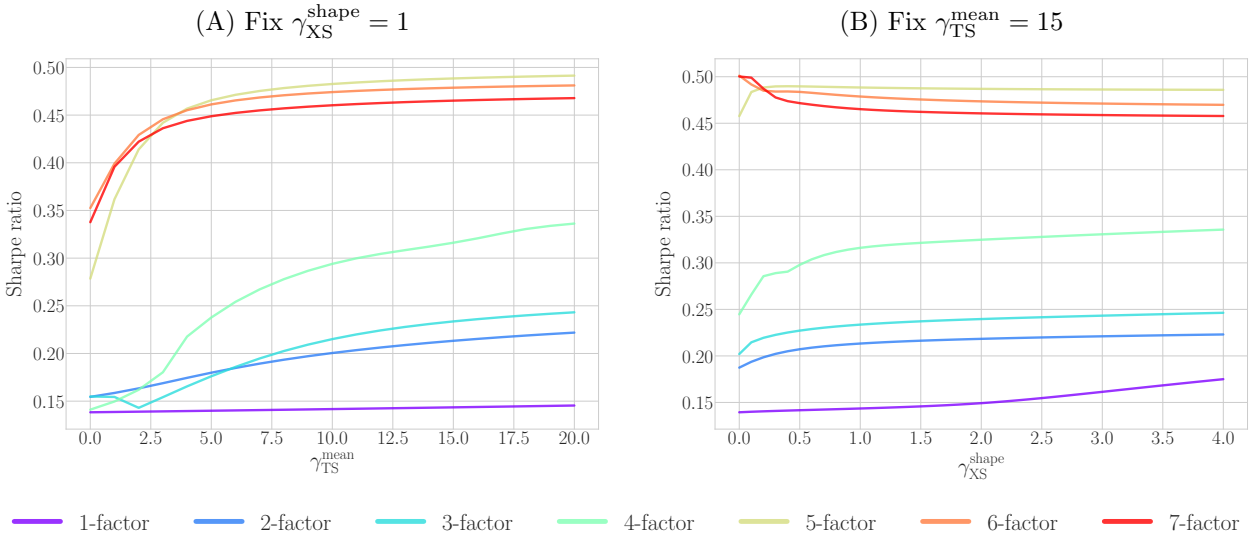
3.4 Combining cross-sectional and time-series targets

We show that combining multiple targets improves asset pricing models. The best factor models take advantage of multiple signals provided by the cross-sectional and time-series constraints. The different penalties are partial complements, but also partial substitutes. After having studied the cross-sectional and time-series targets separately, we now investigate models that exploit both types of targets simultaneously. We start our analysis with combining the cross-sectional shape and time-series mean targets.

Figure 11 shows the out-of-sample Sharpe ratios of the cross-sectional time-series target PCA models that combine the shape and mean targets. We fix one target at its benchmark value, and study the effect of varying the second target for different numbers of factors. Panel A of Figure 11 shows that imposing a mean penalty always improves the Sharpe ratios of the models. Here, we fix the shape target at its benchmark value $\gamma_{\text{XS}}^{\text{shape}} = 1$, and vary the mean target $\gamma_{\text{TS}}^{\text{mean}}$. For any number of factors, increasing $\gamma_{\text{TS}}^{\text{mean}}$ substantially increases the Sharpe ratio relative to the case of $\gamma_{\text{TS}}^{\text{mean}} = 0$. This is evidence that the mean target incorporates information that is complementary to the shape target. Panel B of Figure 11 shows that the cross-sectional shape target can also help to increase the Sharpe ratios of time-series mean PCA models with up to five factors. We set the mean target to its benchmark value $\gamma_{\text{TS}}^{\text{mean}} = 15$, and vary $\gamma_{\text{XS}}^{\text{shape}}$. For one to five factors, the models with $\gamma_{\text{XS}}^{\text{shape}} = 1$ have higher Sharpe ratios than those with $\gamma_{\text{XS}}^{\text{shape}} = 0$. This shows that the shape target provides complementary information not included in the mean target.

Figure 11: Sharpe ratios for models that combine XS-shape and TS-mean targets

This figure shows the out-of-sample Sharpe ratios of the XS-TS-target-PCA models that combine the cross-sectional shape and time-series mean targets for different numbers of factors and penalty parameters. Panel A varies $\gamma_{\text{TS}}^{\text{mean}}$ for a fixed benchmark shape target $\gamma_{\text{XS}}^{\text{shape}} = 1$, whereas Panel B varies $\gamma_{\text{XS}}^{\text{shape}}$ for a fixed benchmark mean target $\gamma_{\text{TS}}^{\text{mean}} = 15$.

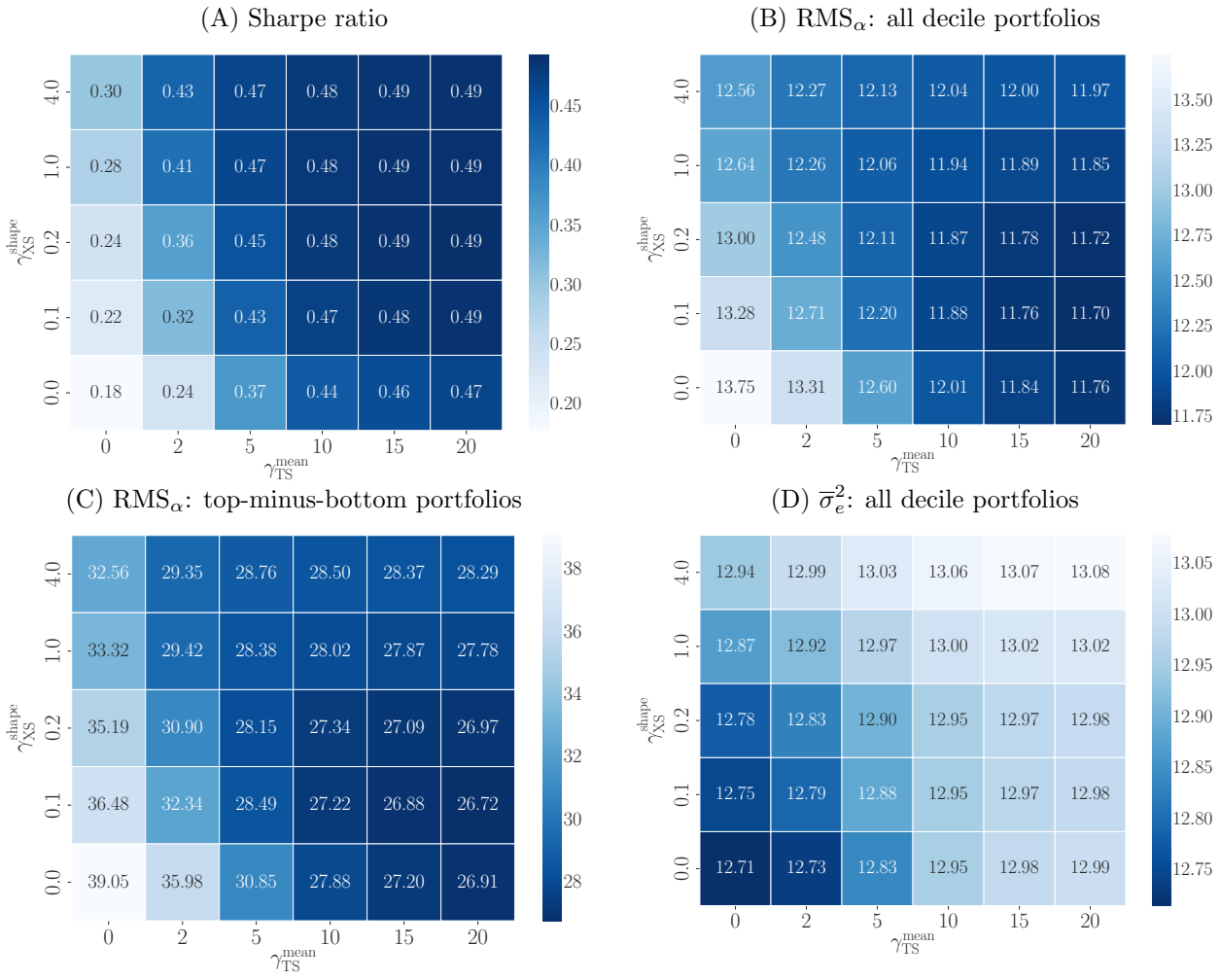


Both targets have economic value. The shape target rewards explaining the spread of high-minus-low anomaly returns, while the mean target rewards explaining the average mean returns of the full cross-section. These are different objectives that identify different weak signals. Adding the shape target to the mean target helps to incorporate the pricing signal of spread portfolios into lower-order factors. In other words, combining both economic constraints provides a better parsimonious factor model.

Figure 12 provides a complete out-of-sample analysis for the benchmark five-factor model. We report the out-of-sample Sharpe ratio, RMS_α, and unexplained variance for different combinations of $\gamma_{\text{XS}}^{\text{shape}}$ and $\gamma_{\text{TS}}^{\text{mean}}$. Panel A demonstrates that we achieve the highest Sharpe ratio of 0.49 when using the benchmark values for both target penalties. Starting with conventional PCA ($\gamma_{\text{XS}}^{\text{shape}} = 0$, $\gamma_{\text{TS}}^{\text{mean}} = 0$) we see that increasing either $\gamma_{\text{XS}}^{\text{shape}}$ or $\gamma_{\text{TS}}^{\text{mean}}$ increases the Sharpe ratio. Hence, both targets contain risk premium information not captured by PCA. For $\gamma_{\text{XS}}^{\text{shape}} = 1$, an increase of $\gamma_{\text{TS}}^{\text{mean}}$ from 0 to 2 has a substantially larger effect than for $\gamma_{\text{XS}}^{\text{shape}} = 0$. This is consistent with the observation that both penalties are partial substitutes. For a high shape target, the model already focuses on specific signals and requires less nudging with the mean target. Similarly, for $\gamma_{\text{TS}}^{\text{mean}} = 15$, the increase in $\gamma_{\text{XS}}^{\text{shape}}$ has a smaller effect than for $\gamma_{\text{TS}}^{\text{mean}} = 0$. This emphasizes again that the signals are partial substitutes. Importantly, the best model combines both signals, which shows that they are also partial complements, that is, each target includes a signal not subsumed by the other target. This complementarity is also reflected in the pricing errors in Panels B and C. The mean penalty improves the pricing errors for the shape target for all decile portfolios as well as for the top-minus-bottom portfolios. Panel D confirms that the explained variation is largely

Figure 12: Performance of five-factor XS-shape-TS-mean-PCA models

This figure shows the out-of-sample performance of XS-shape-TS-mean-PCA with $K = 5$ factors for different cross-sectional shape and time-series mean targets. Panel A reports the maximum Sharpe ratio (SR) of the tangency portfolio. Panel B shows the root-mean-squared pricing error RMS_α for the 370 decile portfolios. Panel C displays the root-mean-squared pricing error RMS_α for the 37 top-minus-bottom portfolios. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



unaffected by both targets, which is consistent with our observation that the targets primarily affect weak factors.

Figure D.9 in the Appendix shows the corresponding results for one to six factors. Our general takeaways carry over to different numbers of factors. Interestingly, for a four-factor model with $\gamma_{\text{TS}}^{\text{mean}} = 15$, increasing $\gamma_{\text{XS}}^{\text{shape}}$ from 0 to 1 increases the Sharpe ratio by a third from 0.24 to 0.32. Hence, the additional shape target helps to include weaker signals into lower-order factor models. In other words, combining multiple targets can reward us with a more parsimonious representation that reflects more risk premium information.

In addition to better pricing performance, combining multiple targets affects the composition of the factor structure and can improve its interpretability. Figure C.6 compares the com-

positions of the factors for the combined XS-shape-TS-mean-PCA model with those for the TS-mean-PCA model. Combining the shape target with the mean target leads to a more symmetric long-short structure in the portfolio weights, which can help with their interpretation. Interestingly, XS-shape-TS-mean-PCA emphasizes different characteristic groups. The fourth factor in the combined shape and mean model puts a larger weight on the trading frictions category. This category does not receive a high weight for a pure TS-mean-PCA, but is emphasized in the fourth XS-shape-PCA factor. This change in the composition of the fourth factor explains the better performance for the four-factor model that combines both targets.

For brevity, we defer the complete analysis for XS-shape-TS-alpha-PCA, which combines the shape target with the FF5-alpha target, to the Internet Appendix. The qualitative results and insights are similar to combining the shape with the mean target. The best model combines both, the shape and time-series alpha targets. We observe the strongest effect for a four-factor model. We conclude that the shape and alpha targets are partial substitutes and partial complements. An optimal model uses both to provide a parsimonious factor model representation.

Table 2 provides an out-of-sample comparison of all models that we have considered so far. We summarize the main metrics for three- and five-factor models including Fama-French, conventional PCA, and five different cross-sectional time-series target PCA models that use one target or combine two targets.

First, the results in Table 2 highlight that different targets nudge the models to focus on different signals, and hence improve different objectives. Combining signals leads to superior models as they combine non-redundant information from multiple economic constraints. The highest Sharpe ratio of 0.49 for a five-factor model is achieved by XS-shape-TS-mean-PCA, closely followed by 0.48 for XS-shape-TS-alpha-PCA. The lowest RMS_α for all 370 deciles (for three and five factors) is achieved by TS-mean-PCA, which optimizes exactly this objective. The pricing error of the FF5-alpha-weighted portfolio is minimized by XS-shape-TS-alpha-PCA with five factors. The TS-alpha penalty targets FF5-alphas, but as the FF5-alpha-weighted portfolio is a long-short portfolio, the shape target provides additional information. In terms of explained variation, all PCA-based models perform very similarly, highlighting again that the penalties affect mainly higher-order weak factors. A shape target greatly improves the XS-shape metric of the joint models. This confirms that models with mean or alpha target obtain a clear long-short factor structure when adding the shape target, which improves interpretability and raises Sharpe ratios.

Second, conventional PCA does generally not outperform Fama-French factors in terms of Sharpe ratio or pricing errors. All XS-TS-target-PCA models that include a time-series target substantially outperform conventional PCA and Fama-French factors for all asset pricing metrics. This finding highlights again that it is crucial to include economic constraints to detect weak asset pricing factors.

Our analysis focuses on demonstrating the effect of target penalties and the number of factors on the out-of-sample performance. For conciseness we report the out-of-sample metrics for different choices of parameters. As the effect of the parameters on the performance is stable over

Table 2: Out-of-sample performance summary for XS-TS-target-PCA

This table summarizes the out-of-sample performance of three- and five-factor models including the Fama-French (FF), conventional PCA, and five different cross-sectional times-series target 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). The benchmark penalty parameter on the cross-sectional shape target is $\gamma_{\text{XS}}^{\text{shape}} = 1$ and that on the time-series mean and alpha targets is $\gamma_{\text{TS}} = 15$. For each model we report the maximum 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 the cross-sectional shape and time-series alpha metrics. We calculate the last two metrics in each rolling window and report their time-series averages. 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
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

time, we can form optimal, implementable trading strategies, that is, we can use the first part of the time-series sample to select the penalties and number of factors to use the resulting models for out-of-sample investment and asset pricing on the second part of the sample. We demonstrate this result in the Internet Appendix. We use the first 300 months as training and validation samples, and then evaluate the model on the test sample of the last 350 months. Specifically, we use five-fold cross-validation on the first part of the data to select the penalty parameters to optimize a specific criterion. This is similar to the analysis in Lettau and Pelger (2020b) and Kozak et al. (2020). Given the optimal choice of penalty parameters, we show that the selected model is among the best-performing models on the test sample.

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 macroeconomic state variables that affect the marginal utility of consumption of a representative investor should be priced. To incorporate this economic restriction into the recovery of factor models, we consider the cross-sectional target matrix Λ^{XS} defined in (7), whose columns contain the covariances between asset

returns and seven macro factor innovations as described below. Thus, the cross-sectional macro target nudges the latent factors to span the same space as the Fama-MacBeth-type mimicking portfolios of the macroeconomic innovations, which are the managed target portfolios in this case.

To construct the cross-sectional macro target, we use monthly data for 127 macroeconomic variables from FRED-MD. We construct the innovations to these macroeconomic variables using the following procedure. First, we transform these variables with 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.[7] Finally, 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.

At its core, the cross-sectional macro target nudges factors to explain the correlation with macro mimicking portfolios. Those portfolios might or might not carry high risk premia. Hence, the macro target does not necessarily increase the Sharpe ratio of the tangency portfolio if those macro mimicking portfolios are not priced. Our method provides a disciplined way to test this conjecture.

Figure 13 shows the out-of-sample performance of the cross-sectional macro PCA (XS-macro-PCA) models for different numbers of factors and values of the penalty parameter. Panels A and B demonstrate that the cross-sectional macro target helps to increase the Sharpe ratios of low-dimensional models with up to five factors. However, the macro target can decrease the Sharpe ratio for models with six or seven factors. Thus, in general, the macro target is not as effective as the shape, mean, and alpha targets at capturing risk premium information of the cross section.

On the other hand, Panel C in Figure 13 shows that the cross-sectional macro target helps to monotonically decrease the root-mean-squared pricing error for the seven macro mimicking portfolios. For instance, the root-mean-squared pricing error for the five-factor model decreases by over 40% from 9.24 to 5.20 basis points when the penalty increases from zero to 0.5. Thus, consistent with its objective, the macro target does help to price the macro mimicking portfolios better. This enhanced pricing ability does not come at the cost of lower time-series explanatory power because Panel D shows that the average unexplained variance on all decile portfolios does not vary substantially with the macro target.

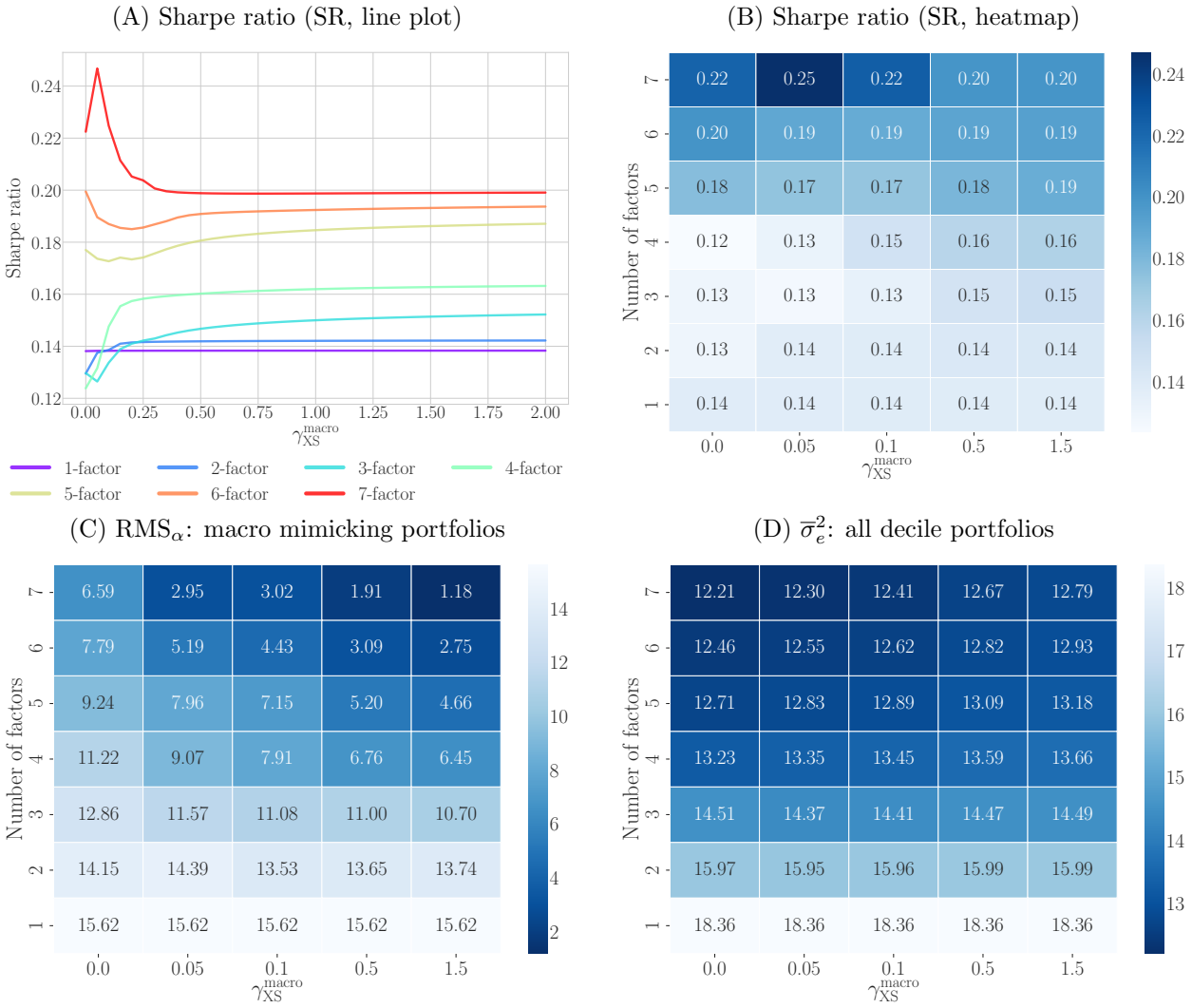
We now consider models that exploit the cross-sectional macro and time-series mean targets simultaneously. Figure 14 shows the out-of-sample Sharpe ratios of the cross-sectional time-series target PCA models that combine the macro and mean targets. We fix one target at its benchmark value, and study the effect of varying the second target for different numbers of factors.

Panel A of Figure 14 shows that imposing a mean penalty always improves the Sharpe ratios of the models. Here, we fix the shape target at its benchmark value $\gamma_{\text{XS}}^{\text{macro}} = 0.5$, and vary the

[7] 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.

Figure 13: Out-of-sample performance of cross-sectional macro PCA models

This figure shows the out-of-sample performance of the cross-sectional macro PCA (XS-macro-PCA) for different numbers of factors and different values of the penalty parameter $\gamma_{\text{XS}}^{\text{macro}}$. Panel A and Panel B report the maximum Sharpe ratio (SR) of the tangency portfolio for up to seven factors. Panel C displays the root-mean-squared pricing error RMS_α of seven macro mimicking portfolios. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.

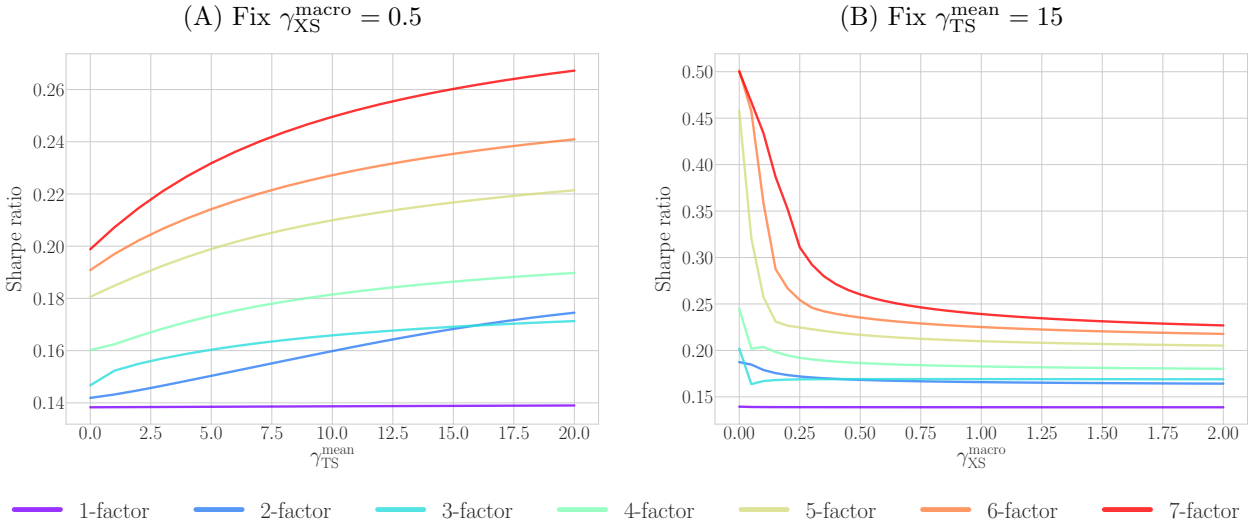


mean target $\gamma_{\text{TS}}^{\text{mean}}$. For any number of factors, increasing $\gamma_{\text{TS}}^{\text{mean}}$ substantially increases the Sharpe ratio relative to the case of $\gamma_{\text{TS}}^{\text{mean}} = 0$. Unlike the cross-sectional macro target, the time-series mean target explicitly increases the signals of factors with high risk premia, and thus helps to identify factors with a higher overall Sharpe ratio.

Panel B of Figure 14, on the other hand, shows that there is a trade-off between pricing the macro mimicking portfolios and the broad cross section of characteristic decile portfolios, suggesting that the macro target fails to capture factors that carry risk premia that explain the expected returns of the entire cross section. We set the mean target to its benchmark value $\gamma_{\text{TS}}^{\text{mean}} = 15$, and

Figure 14: Sharpe ratios for models that combine XS-macro and TS-mean targets

This figure shows the out-of-sample Sharpe ratios of the XS-TS-target-PCA models that combine the cross-sectional macro and time-series mean targets for different numbers of factors and penalty parameters. Panel A varies $\gamma_{\text{TS}}^{\text{mean}}$ for a fixed benchmark macro target $\gamma_{\text{XS}}^{\text{macro}} = 0.5$, whereas Panel B varies $\gamma_{\text{XS}}^{\text{macro}}$ for a fixed benchmark mean target $\gamma_{\text{TS}}^{\text{mean}} = 15$.



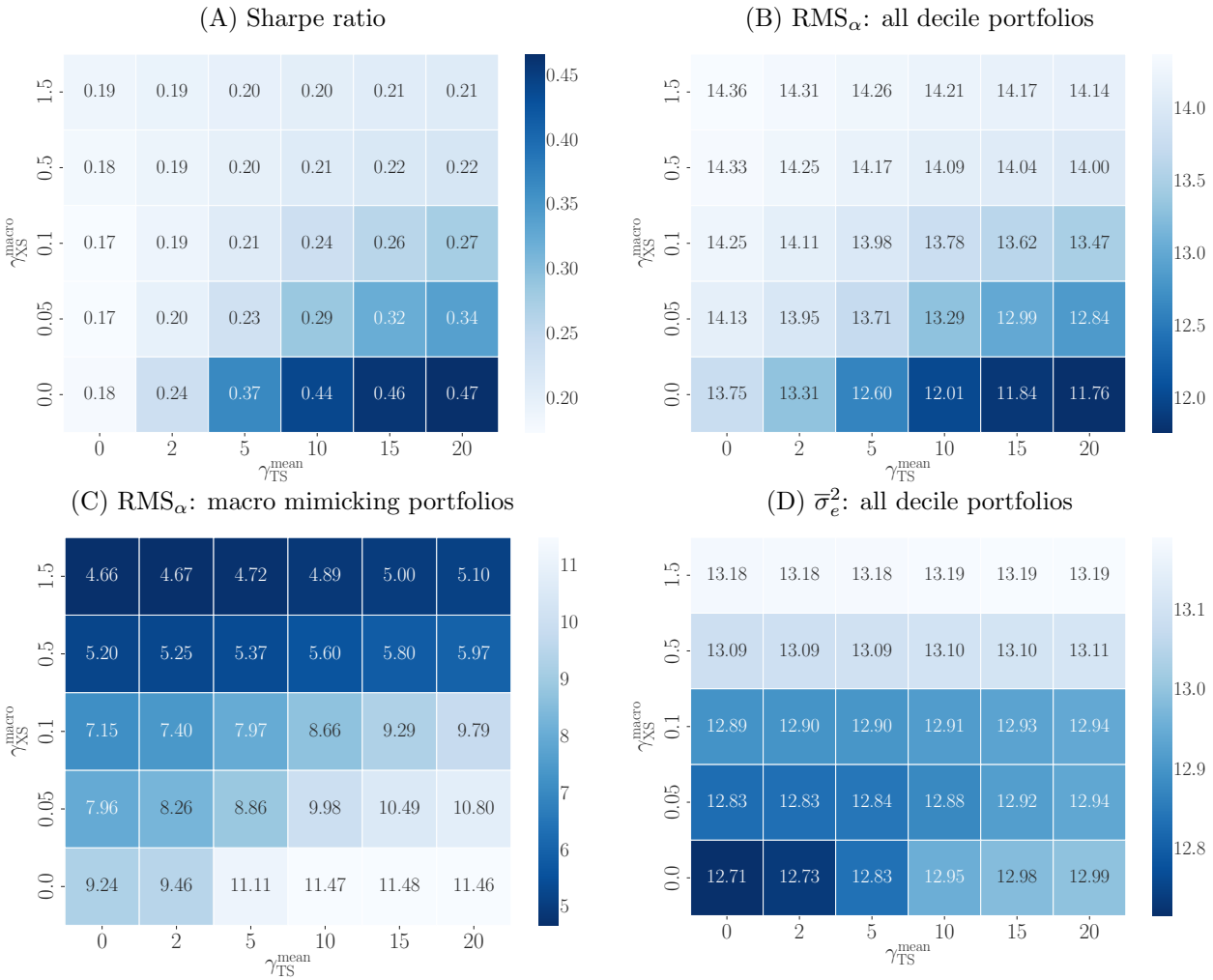
vary $\gamma_{\text{XS}}^{\text{macro}}$. When we increase the penalty on the macro target and thus nudge the factors to span the macro mimicking portfolios, the Sharpe ratios of the time-series mean PCA models decrease substantially. In other words, although Panel C in Figure 13 shows that the macro target helps to identify factors that are better at pricing the macro mimicking portfolios, it also nudges the time-series mean PCA factors in a direction that worsens their performance at pricing a general cross section of assets.

The trade-off between pricing the macro mimicking portfolios and the broad cross section of characteristic decile portfolios is also evident in Figure 15, which shows the out-of-sample performance of the benchmark five-factor model for different penalty parameters on both the cross-sectional macro and time-series mean targets. In addition to Sharpe ratio, we report the root-mean-squared pricing errors for all decile portfolios and for the seven macro mimicking portfolios, and the unexplained variance for the full cross-section.

We find that, for a given positive 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 all decile portfolios, but it decreases the pricing error for the seven macro mimicking portfolios. This is consistent with Panel B of Figure 14 and suggests that the covariances between asset returns and macroeconomic innovations cannot generate the return spread of the large cross-section of characteristic decile portfolios. Panel C in Figure 15 also demonstrates the flip-side, namely that increasing the penalty on the mean target increases the root-mean-squared pricing error for the seven macro mimicking portfolios, but decreases the RMS_α for the full cross-section. In summary,

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

This figure shows the out-of-sample performance of XS-macro-TS-mean-PCA with $K = 5$ factors for different cross-sectional macro and time-series mean targets. Panel A reports the maximum Sharpe ratio (SR) of the tangency portfolio. Panel B shows the root-mean-squared pricing error RMS_α for the 370 decile portfolios. Panel C displays the root-mean-squared pricing error RMS_α for seven macro mimicking portfolios. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.



there is a trade-off between pricing the macro mimicking portfolios or the full cross-section, which is reflected in the effect of the two target penalties.

To summarize this section, Table 3 compares the performance of the Fama-French and conventional PCA models with that of three 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). Table 3 shows that each target helps to improve out-of-sample performance in terms of the particular criterion it is designed to address. Also, although there is a trade-off between pricing the macro mimicking portfolios and pricing the entire cross-section of characteristic decile portfolios, the goals of the cross-sectional

Table 3: Out-of-sample performance of summary for XS-TS-target-PCA with macro target

This table summarizes the out-of-sample performance of three- and five-factor models including the Fama-French (FF), conventional PCA, XS-macro-PCA, XS-macro-TS-mean-PCA, and XS-macro-TS-alpha-PCA. The benchmark penalty parameter on the cross-sectional macro target is $\gamma_{\text{XS}}^{\text{macro}} = 0.5$ and that on the time-series mean and alpha targets is $\gamma_{\text{TS}} = 15$. For each model we report the maximum Sharpe ratio, root-mean-squared pricing error across the 370 decile portfolios, root-mean-squared pricing error for the seven macro mimicking portfolios, average unexplained variances across the 370 decile portfolios, and the cross-sectional macro and time-series alpha metrics. We calculate the last two metrics in each rolling window and report their time-series averages. 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
Three-factor models							
FF	0.16	17.23	18.05	20.86	16.63	-	-
PCA	0.13	15.56	12.86	22.86	14.51	78.37	39.60
XS-macro-PCA	0.15	15.39	11.00	22.88	14.47	**99.29**	35.62
XS-macro-TS-mean-PCA	**0.17**	**14.86**	9.98	22.32	14.47	98.73	39.29
XS-macro-TS-alpha-PCA	**0.17**	15.11	**8.95**	**20.13**	**14.39**	99.05	**44.92**
Five-factor models							
FF	**0.31**	**13.66**	11.91	28.41	15.64	-	-
PCA	0.18	13.75	9.24	20.45	**12.71**	65.41	45.29
XS-macro-PCA	0.18	14.33	5.20	20.03	13.09	**98.10**	41.54
XS-macro-TS-mean-PCA	0.22	14.04	5.80	19.24	13.10	96.82	45.98
XS-macro-TS-alpha-PCA	0.21	14.40	**4.08**	**17.62**	13.10	97.57	**54.51**

and time-series targets can be achieved simultaneously in some cases, and thus, assigning positive penalties to both targets can help to recover factors that achieve certain objectives. For example, the XS-macro-TS-alpha-PCA models deliver the lowest out-of-sample root-mean-squared pricing errors for the macro mimicking portfolios and the FF5-alpha-weighted portfolio. In conclusion, we provide a framework to estimate factors that are optimal for a specific economic objective and to evaluate in a disciplined way if this objective provides additional pricing information.

4 Conclusion

This paper addresses the problem of identifying a parsimonious factor model to price a large cross section of asset returns. The challenge is both statistical and economical—returns are noisy, signals are weak, and economic structure can be violated by purely statistical models. Thus, we propose a general framework to combine statistical learning with economic targets. The method is data-driven but guided by economic insight, enabling researchers to encode their beliefs about the economic properties of risk factors. This allows us to detect weak factors that improve cross-sectional asset pricing.

Our framework accommodates general time-series and cross-sectional targets and, in our empirical analysis based on a high-dimensional cross section of assets, we consider four types of targets. First, shape restrictions related to the patterns of asset loadings on risk factors (empirically, we focus on the monotonicity of expected returns in decile-sorted portfolios). Second, a spanning

target, implying that recovered latent risk factors should (at least) reflect the pricing ability of candidate exogenous state variables such as macroeconomic innovations. Third, an APT-implied restriction that systematic risk factors should span the vector of expected returns (not just the sources of time-series variation). Fourth, an explicit nudge of asset-pricing factors to span not only the main sources of time-series variation, often captured by a few popular reduced-form factors, but also the pricing errors generated by the latter, with the aim of better spanning the underlying SDF and cross-section of asset returns.

We find that imposing informative economic targets 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 generate a higher out-of-sample Sharpe ratio and yield lower pricing errors compared to the traditional approach. We contribute to a 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 asset-pricing targets 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 targets consistent with both structural and reduced-form insights about the cross-section of asset returns and risk factors that drive it. Thus, we provide a disciplined approach to evaluate the usefulness of different economic targets. Lastly, our method can also serve as a building block to enhance other methods by replacing conventional PCA factors with economically constrained latent factors.

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, S., M. Pelger, and J. Zhu, 2019, Forest through the trees: Building cross-sections of stock returns, forthcoming in *Journal of Finance*.

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

Chamberlain, Gary, and Michael Rothschild, 1983, Arbitrage, factor structure, and mean-variance analysis on large asset markets, *Econometrica: Journal of the Econometric Society* 51, 1281–1304.

Chen, Luyang, Markus Pelger, and Jason Zhu, 2024a, Deep learning in asset pricing, *Management Science* 70, 714–750.

Chen, Qihui, Nikolai L. Roussanov, and Xiaoliang Wang, 2023, Semiparametric conditional factor models: Estimation and inference, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3984633.

Chen, Zhongtian, Nikolai L. Roussanov, Xiaoliang Wang, and Dongchen Zou, 2024b, Common risk factors in the returns on stocks, bonds (and options), redux, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4703281.

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66, 1047–1108.

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.

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.

DeMiguel, Victor, Javier Gil-Bazo, Francisco J. Nogales, and André A.P. Santos, 2023, Machine learning and fund characteristics help to select mutual funds with positive alpha, *Journal of Financial Economics* 150, 103737.

Duan, Junting, Markus Pelger, and Ruoxuan Xiong, 2023, Target PCA: Transfer learning large dimensional panel data, *Journal of Econometrics* 105521.

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.

Filipovic, D., M. Pelger, and Y. Ye, 2022a, Shrinking the term structure, available at https://ssrn.com/abstract=4182649.

Filipovic, D., M. Pelger, and Y. Ye, 2022b, Stripping the discount curve – A robust machine learning approach Available at SSRN: https://ssrn.com/abstract=4058150.

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

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

Giglio, Stefano, Dacheng Xiu, and Dake Zhang, 2022, Test assets and weak factors, available at https://ssrn.com/abstract=3768081.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu, 2020, Empirical asset pricing via machine learning, *Review of Financial Studies* 33, 2223–2273.

Harvey, C. R., Y. Liu, and H. Zhu, 2016, ... and the cross-section of expected returns., *Review of Financial Studies* 29, 5–68.

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

Huang, Dashan, Fuwei Jiang, Kunpeng Li, Guoshi Tong, and Guofu Zhou, 2021, Scaled pca: A new approach to dimension reduction, *Management Science* 68, 1591–2376.

Jensen, Theis I., Bryan Kelly, and Lasse H. Pedersen, 2023, Is there a replication crisis in finance?, *The Journal of Finance* 78, 2465–2518.

Kaniel, Ron, Zihan Lin, Markus Pelger, and Stijn Van Nieuwerburgh, 2023, Machine-learning the skill of mutual fund managers, *Journal of Financial Economics* 150, 94–138.

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.

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

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.

Li, Bin, and Alberto G. Rossi, 2020, Selecting mutual funds from the stocks they hold: A machine learning approach, available at https://ssrn.com/abstract=3737667.

Li, Sophia Z., and Yushan Tang, 2022, Automated risk forecasting, in *Automated Risk Forecasting*.

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.

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.

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

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.

Zaffaroni, Paolo, 2022, Factor models for conditional asset pricing, Imperial Business School working paper.

A Proof of all Propositions

In this appendix, we provide a proof for Propositions 1–3 in the main body of the manuscript. To do this, we first state a general Proposition A.1 that encompasses Propositions 1–3 and then we provide a proof for Proposition A.1.

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 (13) 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 the matrix whose columns contain the first* K *eigenvectors of matrix (A1), then the estimated factor loadings and factor returns are*

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

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

   (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 the matrix whose columns contain the first* K *eigenvectors of matrix (A4), then the estimated factor returns and factor loadings are*

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

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

2. *In the special case of XS-target-PCA, the latent-factor model that optimizes the XS-target-PCA objective in (5) 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-target-PCA objective in* (10) *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*

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

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*

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

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

Proof. We now prove the second part of Proposition A.1. The proofs for the other two parts are very similar, and thus, we omit them to simplify the exposition. We prove the second part of Proposition A.1 in two steps.

Step 1: First-order condition with respect to F

The first-order condition of problem (5) with respect to F is

$$\begin{aligned} & -\frac{2}{NT}\left(R - F\beta^\top\right)\beta - \frac{2\gamma_{\text{XS}}}{LT}\left(RP_\Lambda^{\text{XS}} - F\beta^\top P_\Lambda^{\text{XS}}\right)P_\Lambda^{\text{XS}}\beta = 0 \\ \Rightarrow \quad & F = R\left(I_N + \frac{\gamma_{\text{XS}}N}{L}P_\Lambda^{\text{XS}}\right)\beta\left[\beta^\top\left(I_N + \frac{\gamma_{\text{XS}}N}{L}P_\Lambda^{\text{XS}}\right)\beta\right]^{-1}. \end{aligned} \tag{A15}$$

Let

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

where $\tilde{\gamma}_{\text{XS}} = \sqrt{\gamma_{\text{XS}}N/L+1} - 1$. Note that problem (5) is equivalent to

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

Plugging equations (A15) and (A16) into (A17), we obtain that problem (5) becomes

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

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

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

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

Step 2: First-order condition with respect to β

The first-order condition of problem (5) with respect to β is

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

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

Plugging equation (A21) into (A17), we obtain that problem (5) 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 buy the top-decile and sell 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 lists the 37 characteristics and classifies them 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 percentage average return 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.

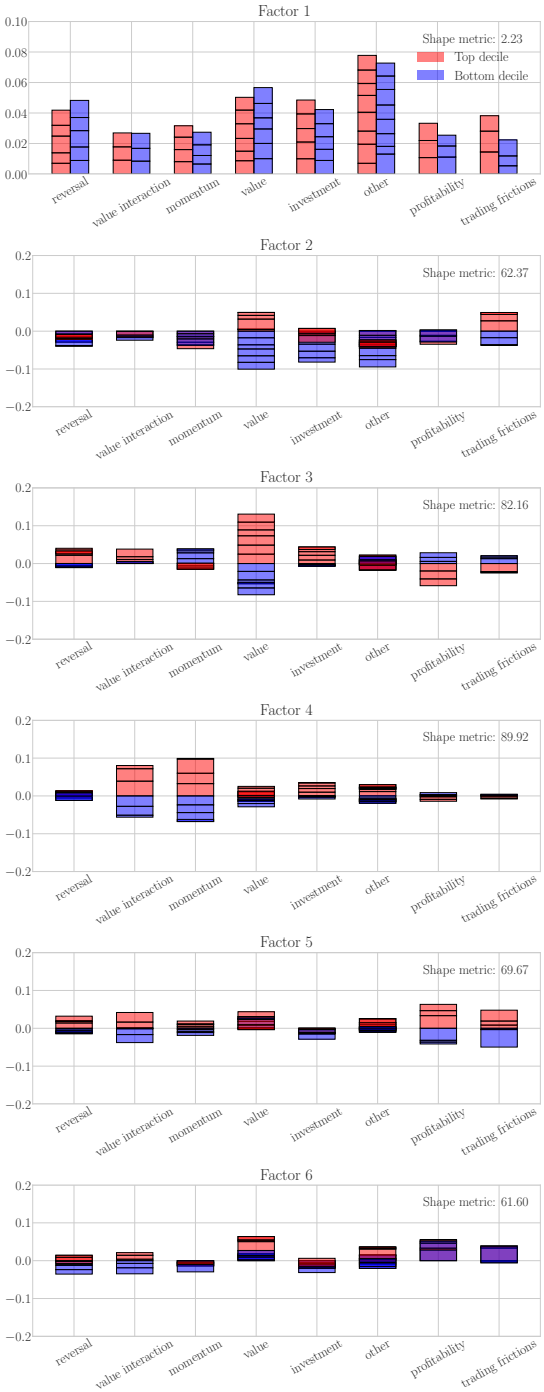
Category	Characteristics	Mean bottom-decile	Mean top-decile	SR top-minus-bottom
reversal	*lrrev, strev, indmomrev, indrrev, indrrevlv*	0.18	1.02	0.23
value interaction	*valmom, valmomprof, valprof*	0.34	1.03	0.16
investment	*inv, invcap, igrowth, growth, noa*	0.39	0.72	0.10
momentum	*mom, mom12, indmom, 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 Factor composition for all models

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

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each of the first six PCA factors and XS-shape-PCA factors with $\gamma_{\text{XS}}^{\text{shape}} = 1$, 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. The factor weights are estimated on the full sample.

(A) PCA



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

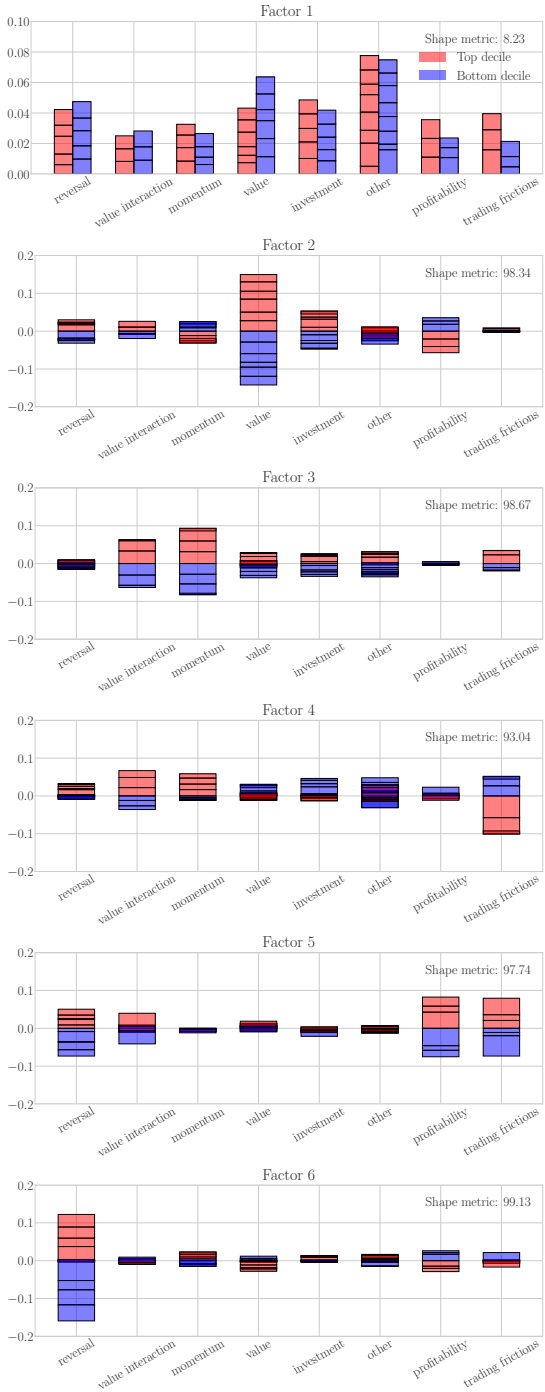


Figure C.2: PCA and XS-shape-PCA factor weights on all decile portfolios

Panels A and B show the portfolio weights on all decile portfolios of the first six PCA factors and XS-shape-PCA factors with $\gamma_{\text{XS}}^{\text{shape}} = 1$, 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). The factor weights are estimated on the full sample.

(A) PCA

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

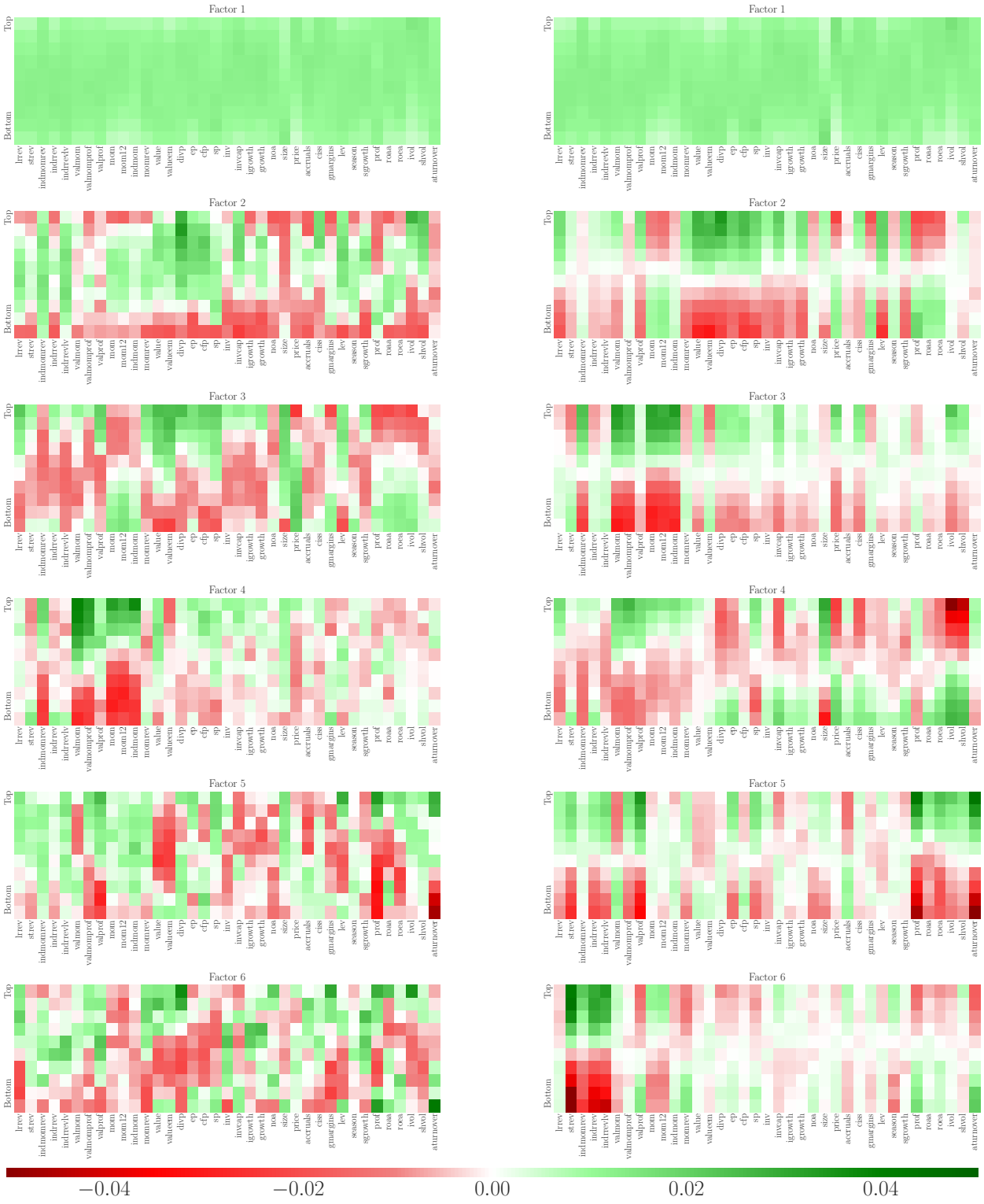


Figure C.3: PCA and TS-alpha-PCA factor weights on extreme-decile portfolios

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each 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. The factor weights are estimated on the full sample.

(A) PCA

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

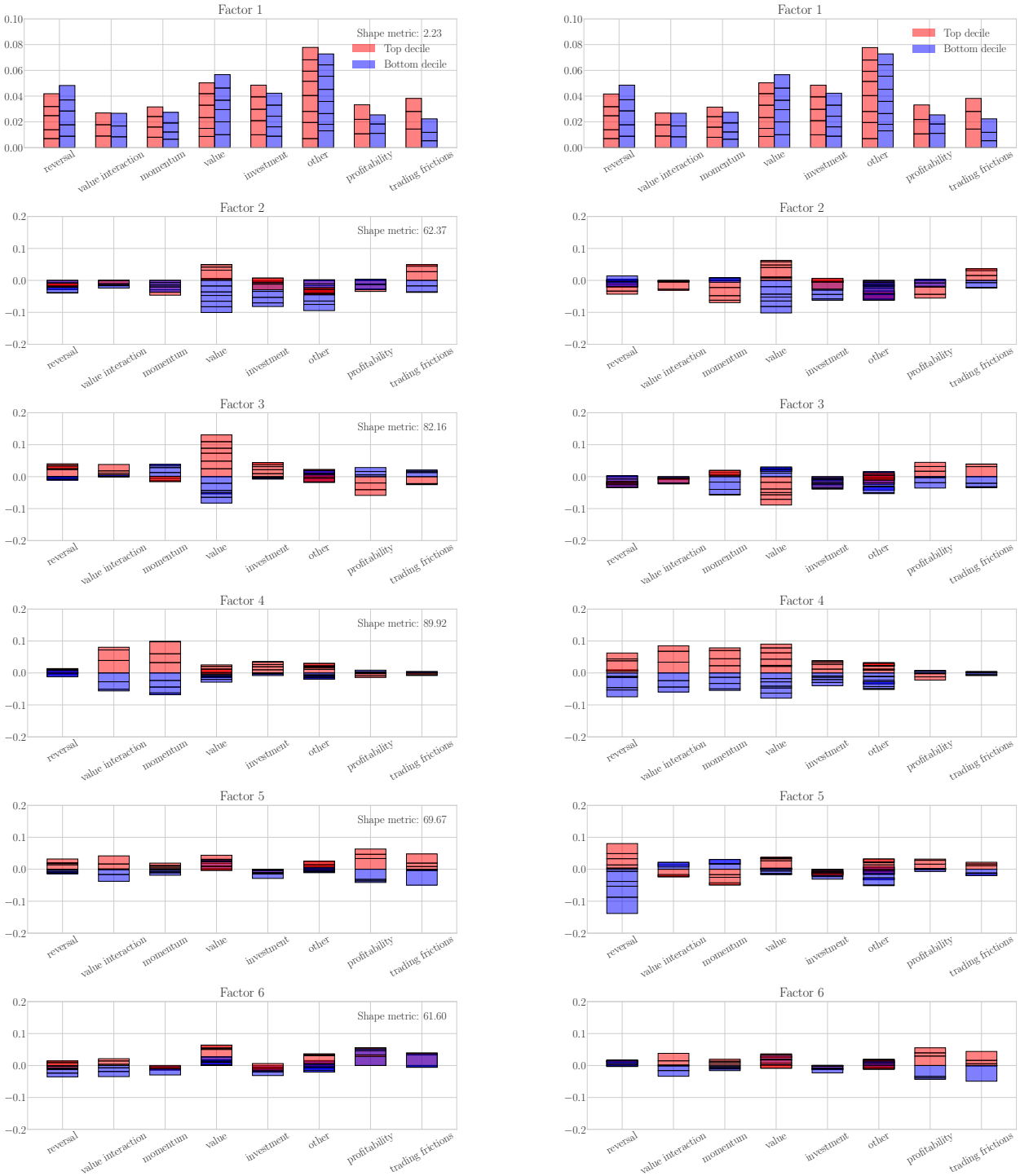


Figure C.4: PCA and TS-mean-PCA factor weights on extreme-decile portfolios

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each 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. The factor weights are estimated on the full sample.

(A) PCA

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

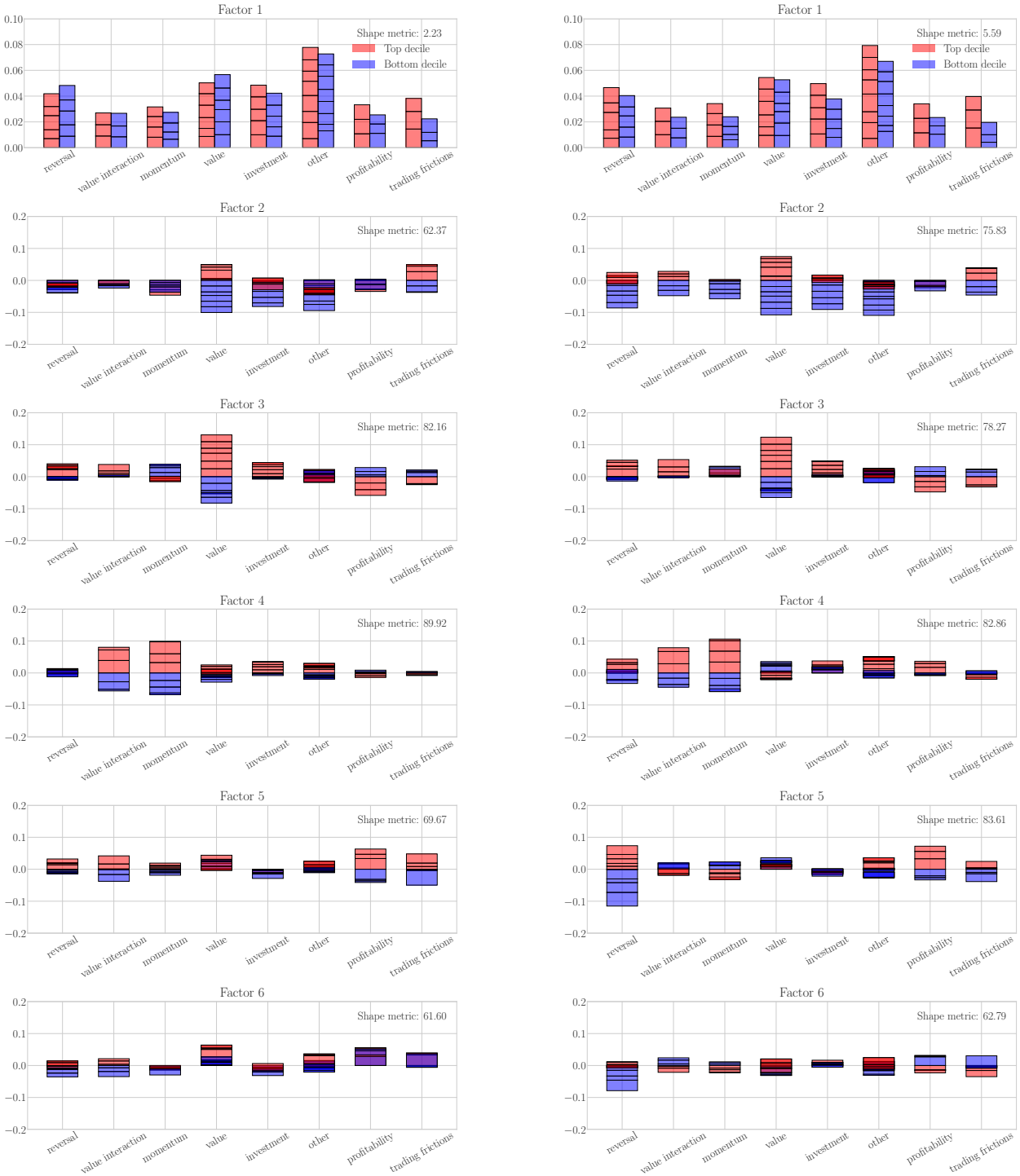
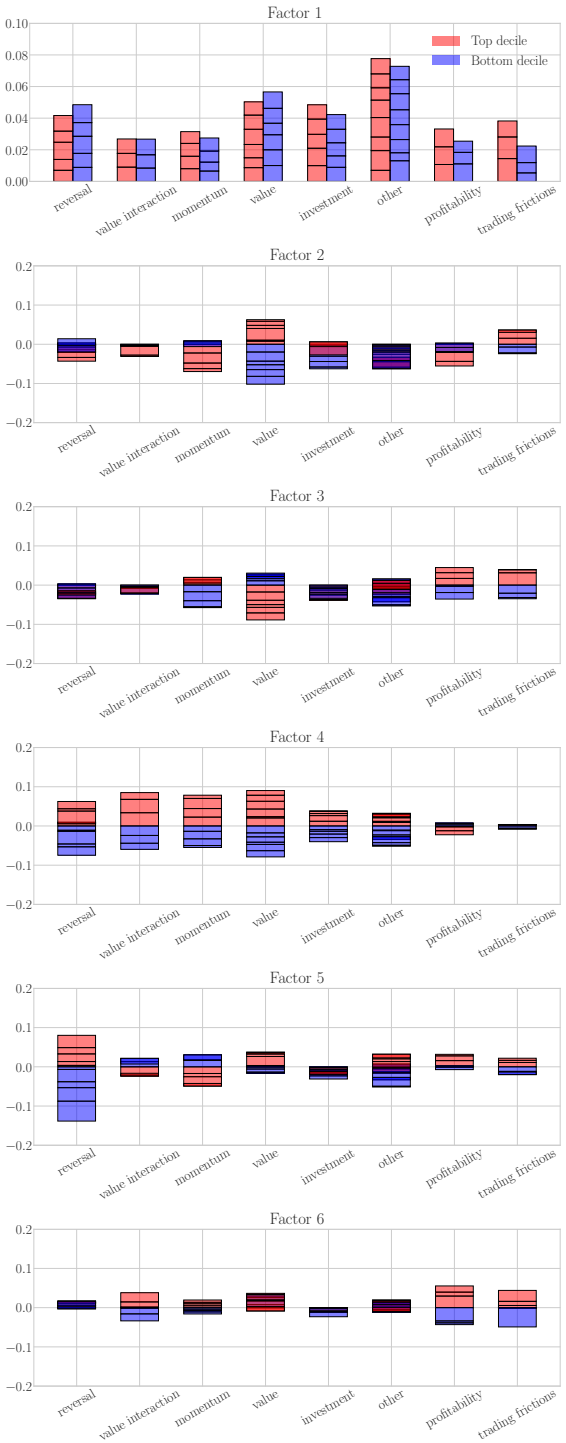


Figure C.5: TS-alpha-PCA and TS-mean-PCA factor weights on extreme-decile portfolios

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each of the first six TS-alpha-PCA factors with $\gamma_{\text{TS}}^{\text{alpha}} = 15$ 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. The factor weights are estimated on the full sample.

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



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

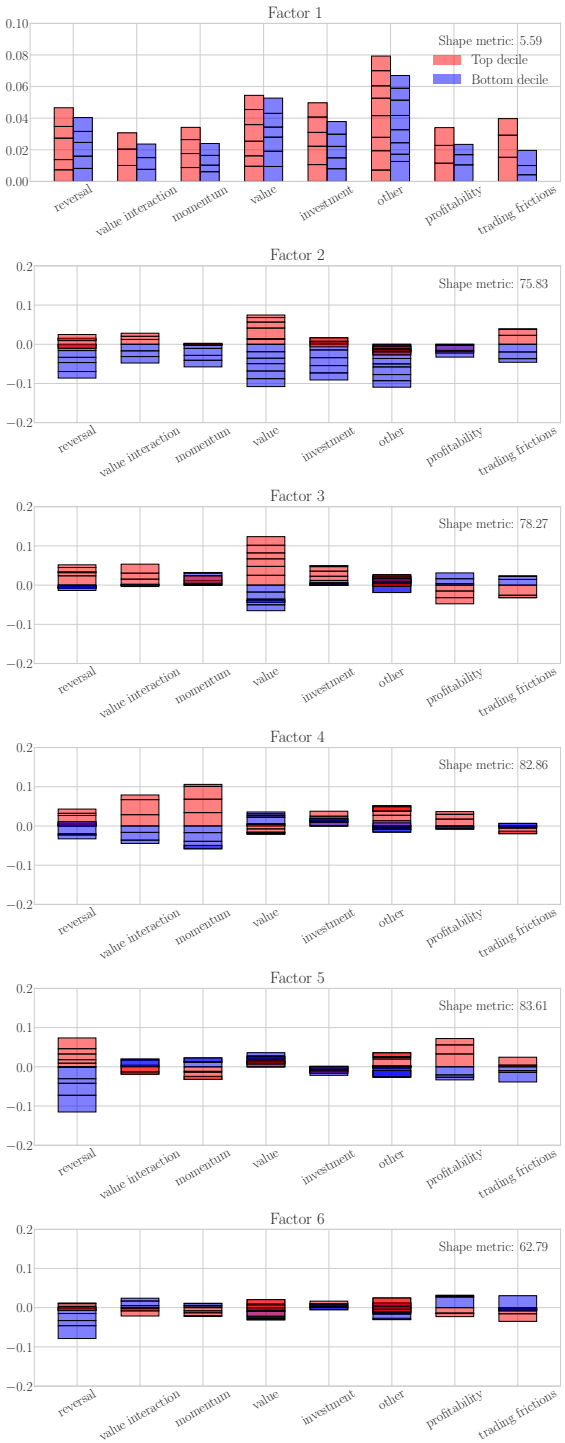
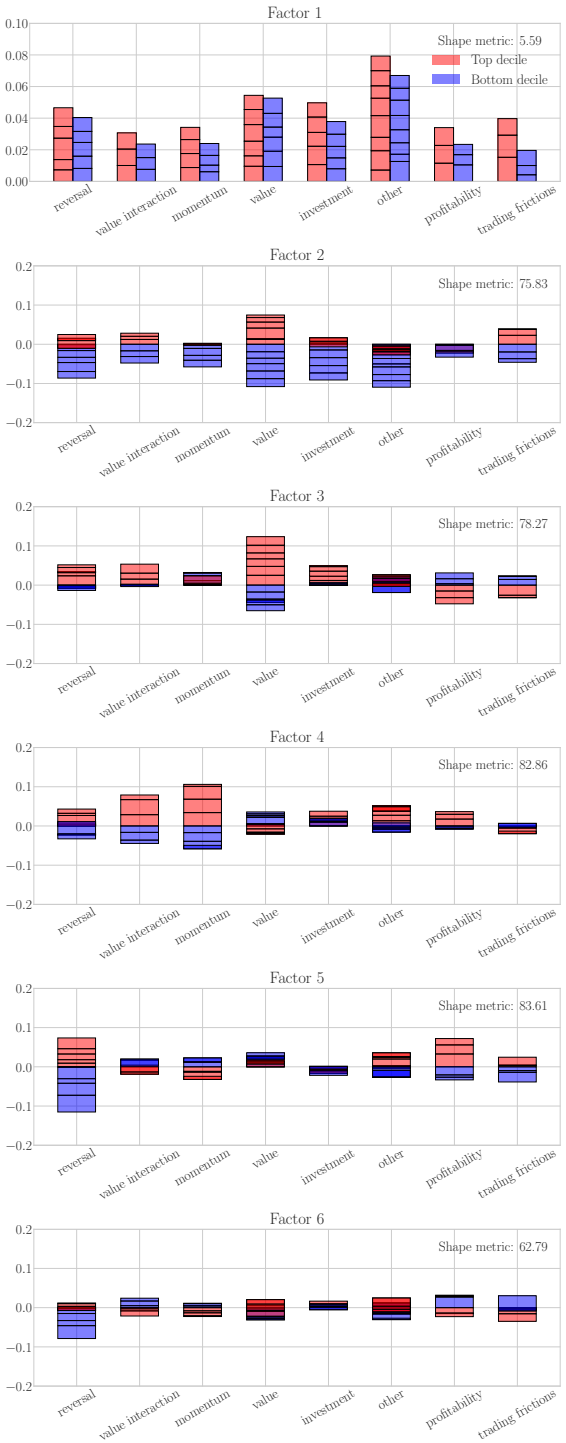


Figure C.6: TS-mean-PCA and XS-shape-TS-mean-PCA factor weights on extreme portfolios

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each of the first six TS-mean-PCA factors with $\gamma_{\text{TS}}^{\text{mean}} = 15$ and XS-shape-TS-mean-PCA factors with $(\gamma_{\text{XS}}^{\text{shape}}, \gamma_{\text{TS}}^{\text{mean}}) = (1, 15)$, 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. The factor weights are estimated on the full sample.

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



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

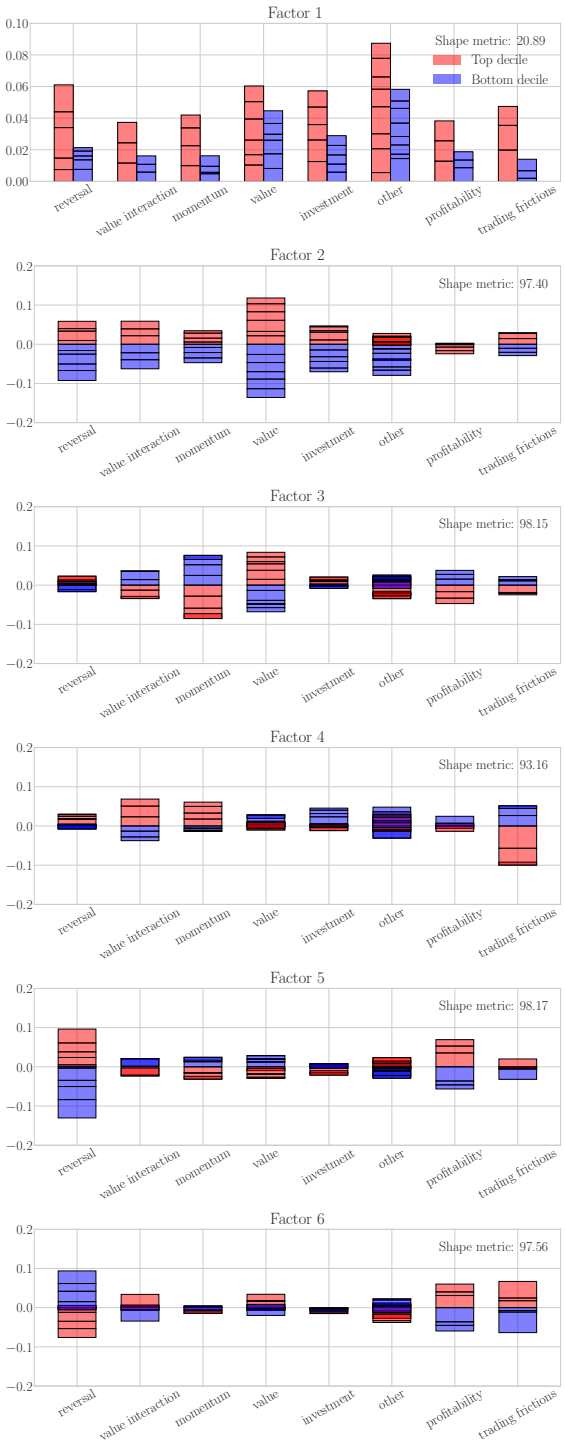
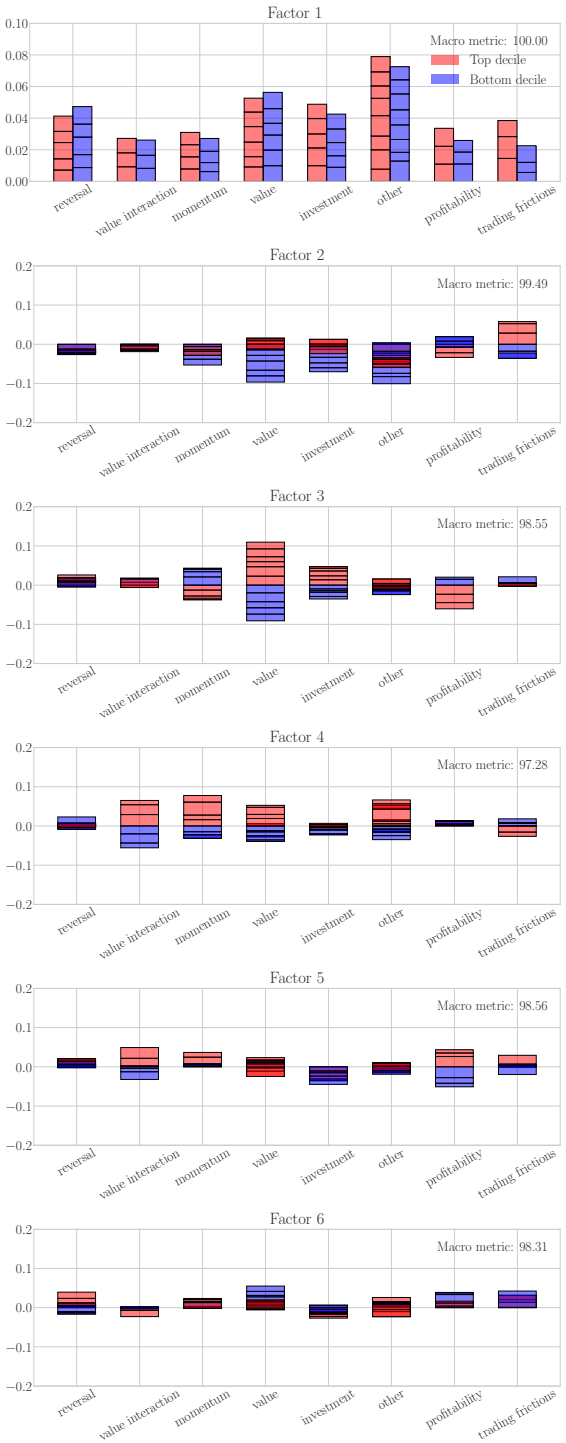


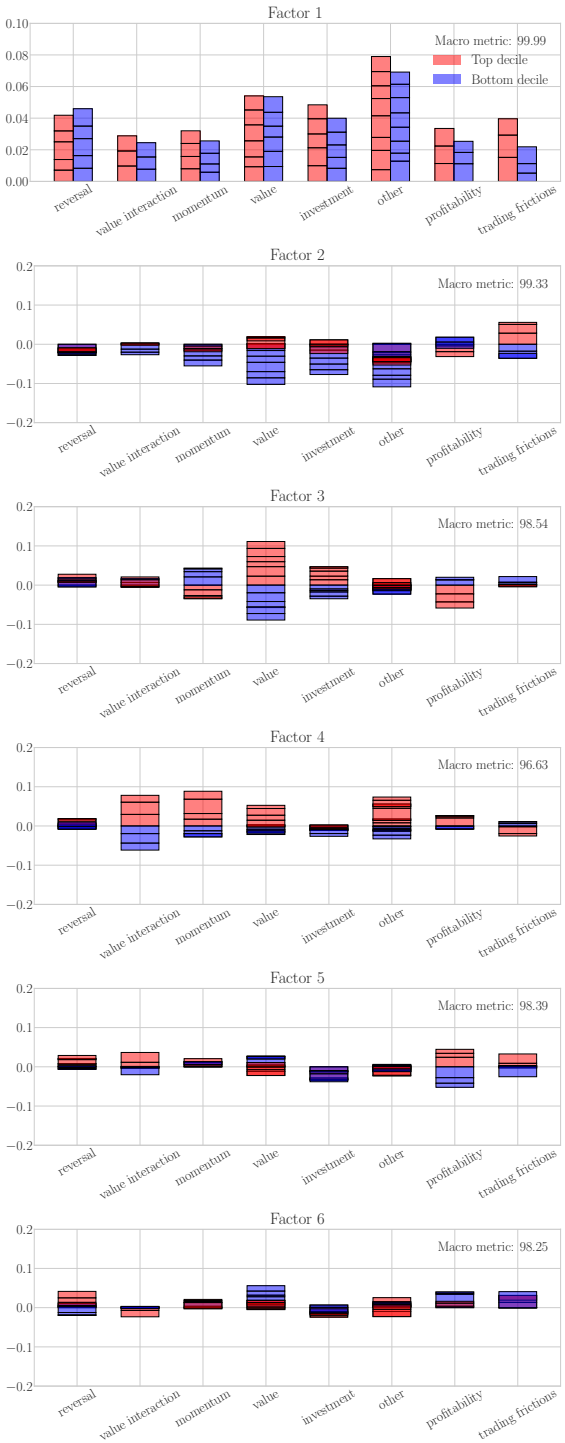
Figure C.7: XS-macro-PCA and XS-macro-TS-mean-PCA factor weights on extreme portfolios

Panels A and B show the portfolio weights on the top and bottom characteristic deciles of each of the first six XS-macro-PCA factors with $\gamma_{\text{XS}}^{\text{macro}} = 0.5$ and XS-macro-TS-mean-PCA factors with $(\gamma_{\text{XS}}^{\text{macro}}, \gamma_{\text{TS}}^{\text{mean}}) = (0.5, 15)$, 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. The factor weights are estimated on the full sample.

(A) XS-macro-PCA with $\gamma_{\text{XS}} = 0.5$



(B) $(\gamma_{\text{XS}}^{\text{macro}}, \gamma_{\text{TS}}^{\text{mean}}) = (0.5, 15)$

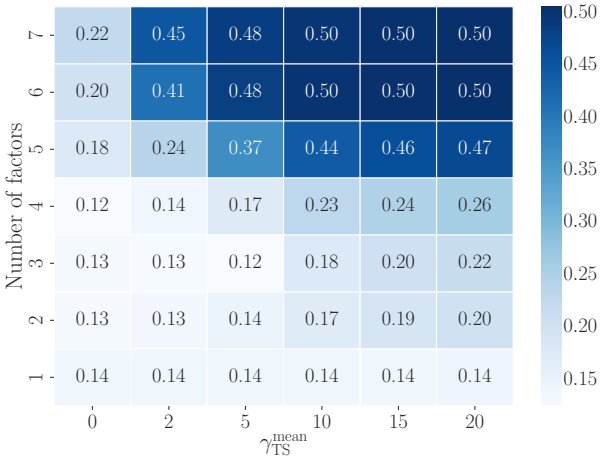


D Results for TS-mean-PCA and XS-shape-TS-mean-PCA

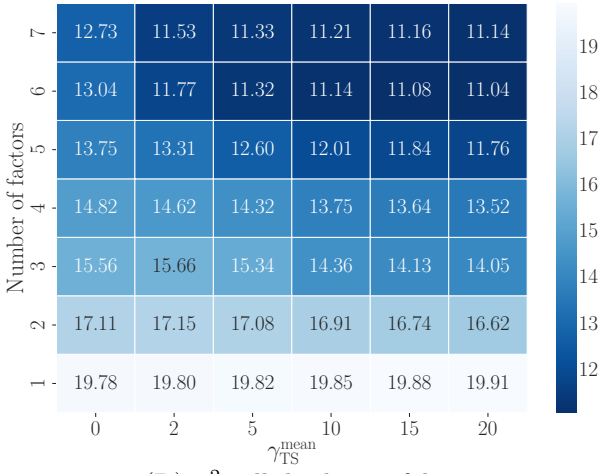
Figure D.8: Performance of time-series mean PCA models

This figure shows the out-of-sample performance of time-series mean PCA (TS-mean-PCA) for different numbers of factors and different values of the penalty parameter $\gamma_{\text{TS}}^{\text{mean}}$. Panel A reports the maximum Sharpe ratio (SR) of the tangency portfolio for up to seven factors. Panel B shows the root-mean-squared pricing error RMS_α for the 370 decile portfolios. Panel C displays the pricing error for the *FF5-alpha-weighted portfolio*, which is a portfolio where the decile weights are proportional to the decile's alpha with respect to FF5. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.

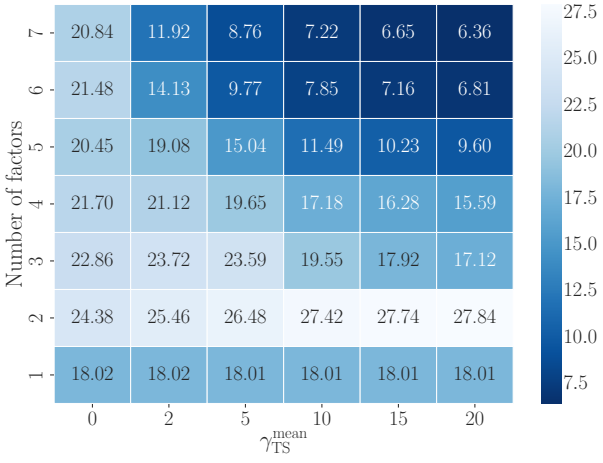
(A) Sharpe ratio



(B) RMS_α: all decile portfolios



(C) Pricing error: FF5-alpha-weighted portfolio



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

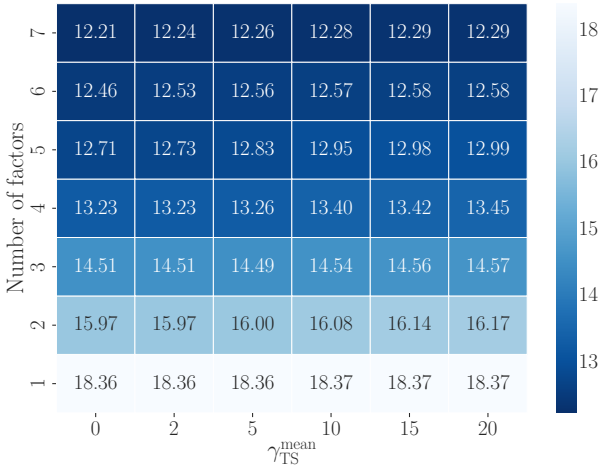


Figure D.9: Performance of XS-shape-TS-mean-PCA models

This figure shows the out-of-sample performance of XS-shape-TS-mean-PCA with $K = 1, \ldots, 6$ factors for different cross-sectional shape and time-series mean targets. Panel A reports the maximum Sharpe ratio (SR) of the tangency portfolio. Panel B shows the root-mean-squared pricing error RMS_α for the 370 decile portfolios. Panel C displays the root-mean-squared pricing error RMS_α for the 37 top-minus-bottom portfolios. Panel D shows the unexplained variance $\overline{\sigma}_e^2$ of the 370 decile portfolios. Root-mean-squared pricing errors are reported in basis points and average unexplained variances in percentage.

(A) Sharpe ratio

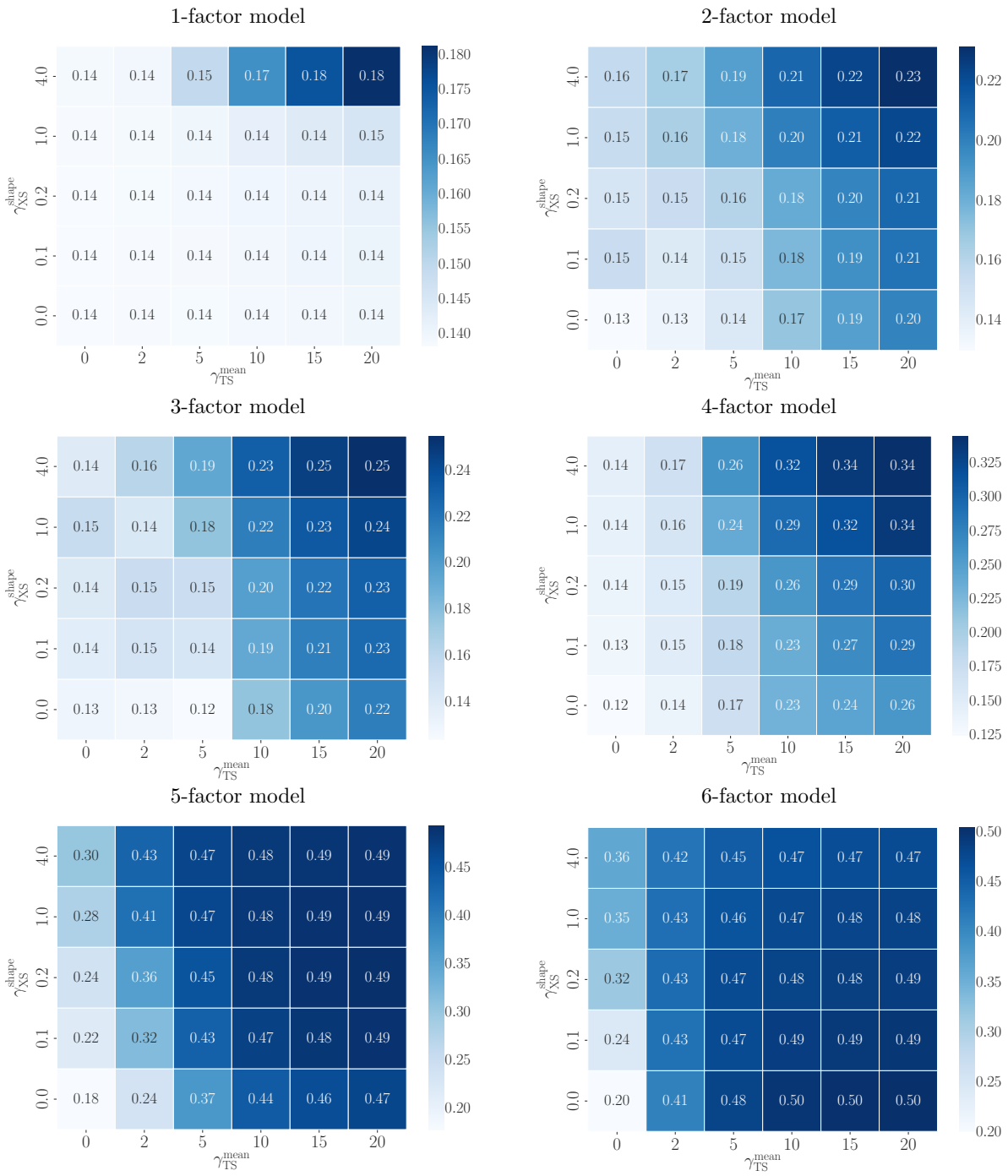


Figure D.9 continued

(B) RMS_α: all decile portfolios

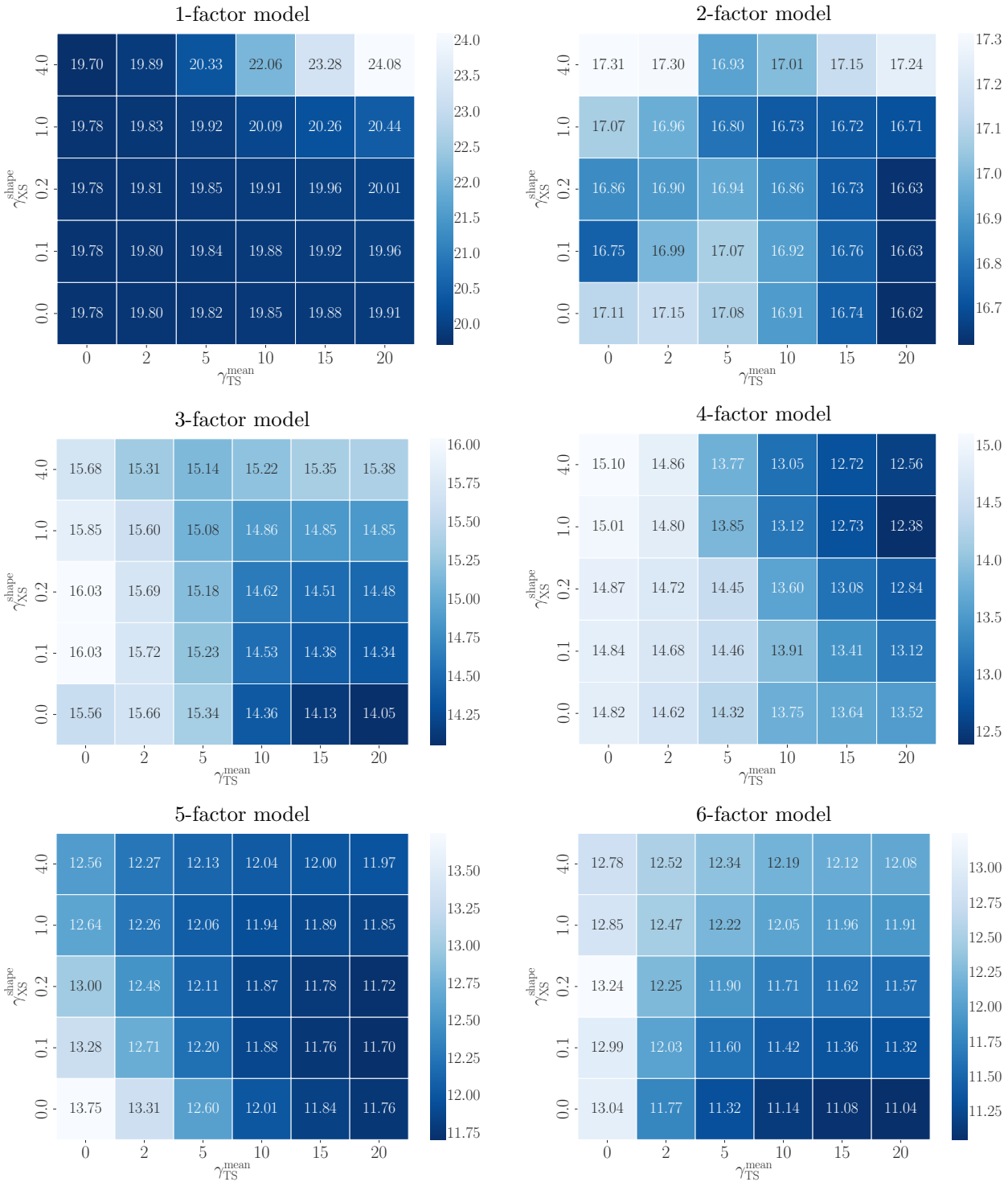


Figure D.9 continued

(C) RMS_α: top-minus-bottom portfolios

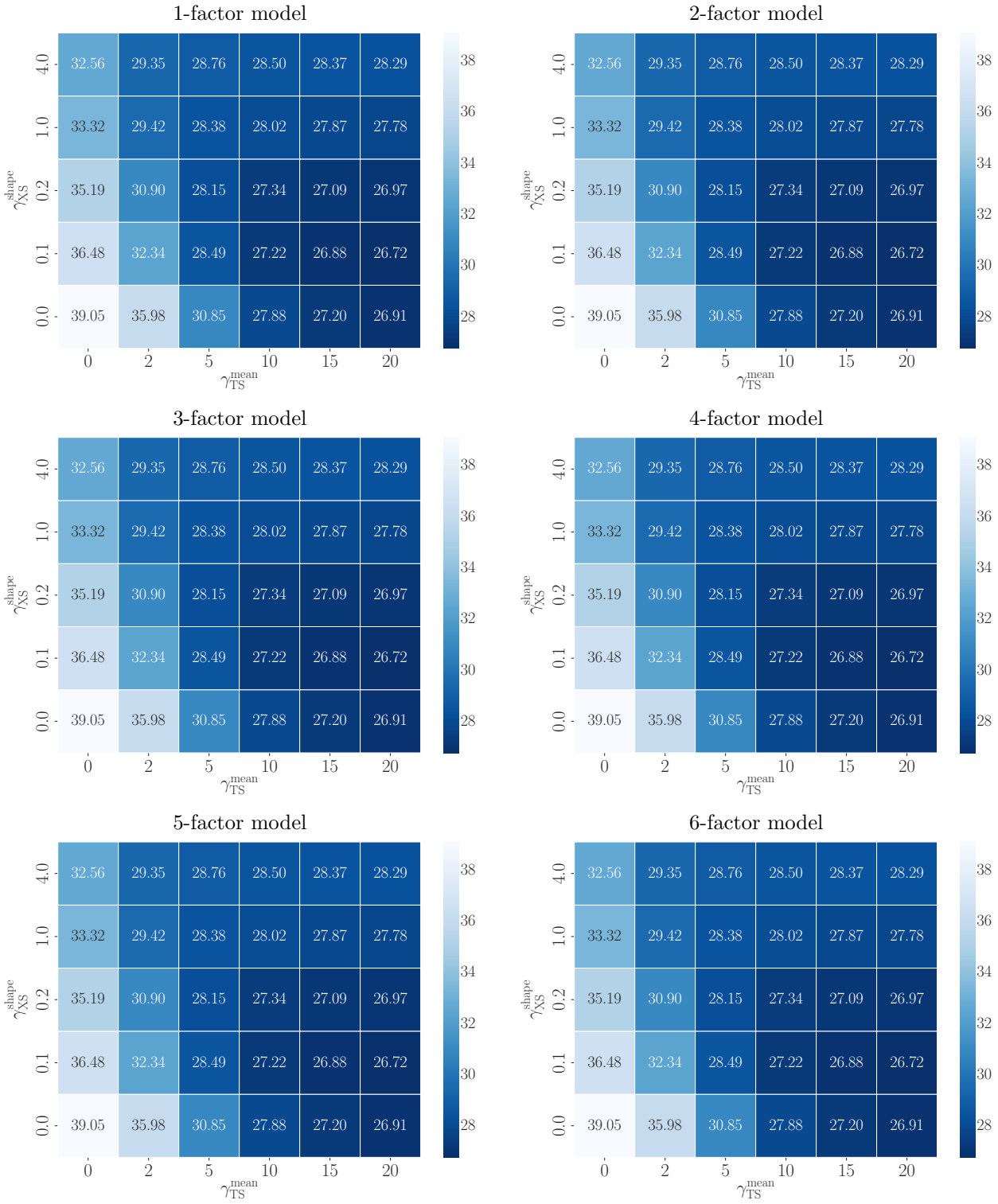


Figure D.9 continued

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

