Moment Fusing: An Informed Construction of Pricing Factors*

Chun-Wei Lin† Ngoc-Khanh Tran‡ Guofu Zhou§

February 13, 2026

Abstract

Existing PCA methods identify common risk factors while ignoring information about expected returns. We propose a general and flexible moment-fusing approach that incorporates risk prices directly into the identification of risk factors. Specifically, we construct a return transformation such that the volatilities of the transformed returns are proportional to the Sharpe ratios of the original assets. We then perform factor analysis on the transformed returns, with or without additional pricing restrictions. Empirically, we find that the moment-fusing factors significantly outperform leading benchmark models in pricing both FX and equity portfolios out-of-sample.

Keywords: Moment Fusing, Pricing Factors, Pricing Restrictions, Sharpe Ratios, Factor Prices.

*We are very grateful to Torben Andersen, Tyler Beason, Jal Bhachech, Ric Colacito, Tuan Dinh, Roger Edelen, Jingzhi Huang, Chris Hrdlicka, Greg Kadlec, Raman Kumar, Anh Le, Greg Leonard, Alex Pecora, Andrea Vedolin, Shixiang Xia, Andreas Stathopoulos, Pengfei Ye and seminar participants at Virginia Tech and Washington University in St. Louis for helpful comments and suggestions. The paper has also benefited from insightful discussions with Thomas Maurer and Brad Paye, to whom we owe a debt of gratitude. All remaining errors are ours.

†Stuart School of Business, Illinois Institute of Technology, Email: clin82@illinoistech.edu.

‡Pamplin College of Business, Virginia Tech, Email: nktran7@vt.edu.

§Olin Business School, Washington University in St. Louis, Email: zhou@wustl.edu.

1 Introduction

That financial assets earn a risk premium as compensation for their exposures to non-diversifiable risks in the market has always been an important guiding principle for asset pricing research. Although the fundamental mean-variance efficiency in the portfolio selection framework Markowitz (1952) conceptually elucidates this risk-return tradeoff, given the sheer number of traded assets, it is challenging to systematically identify important common risk factors that price these asset returns in the cross section and out of sample Cochrane (2011). While the statistical identification of common risk factors concerns mostly the covariance structure of asset returns, and hence, is supported by the covariance-based factor analysis, this prominent analysis framework is oblivious to the first moment (i.e., risk premium) of factors. As a result, identified common risk factors may have insignificant risk premium and factor prices (i.e., Sharpe ratios), giving rise to a poor performance of pricing models constructed from these factors.

The current paper proposes a simple approach to inform the prominent covariance-based analysis with the mean values of asset returns. The approach has two stages and is intuitive. In the first stage, we construct a transformation of asset returns so that the second moments (variances and covariances) of the transformed returns are proportional to the first and higher moments (mean returns and Sharpe ratios) of the original assets. In the second stage, we implement a principal factor analysis (PCA) on the transformed returns, identifying risk factors whose prices are informed by the risk premium and Sharpe ratios of the original assets. The approach preserves the elegance of the PCA methodology. The constructed principal factors are uncorrelated and delineated by their ability to explain the volatilities of the transformed returns, or the Sharpe ratios of the original assets. Since the original and transformed asset returns are equivalent bases of the same return space, the constructed factors pertain to the same fundamental pricing model of interest that governs financial asset returns.

We refer to this intuitive factor construction approach broadly as moment fusing. The moment fusing constructions are flexible and can be extended to incorporate the prominent pricing restrictions recently proposed by Lettau and Pelger (2020). Empirically, we examine and demonstrate the out-of-sample pricing performance of moment-fusing factors and their extended versions (disciplined further by pricing restrictions) in both foreign exchange (FX) and equity markets. While all moment-fusing constructions share the basic feature that means (or other moments) of the original asset returns are fused into the covariance structure of the transformed returns, they differ in their specific fusing objective and procedure. The current paper considers two basic moment-fusing constructions, referred hereafter to as the Sharpe ratio PCA (SR-PCA) and the inverse Sharpe ratio PCA (ISR-PCA), and their combinations and extensions.

The SR-PCA starts with transforming asset returns so that the variances of the transformed returns equal the SRs of the corresponding original returns. A factor analysis implemented on the SR-PCA transformed returns identifies (and prioritizes) factors whose volatilities reflect (and prioritize) high SRs of the original returns. The ISR-PCA starts with tranformating returns so that all transformed mean returns are identical. As a result of this control (i.e., homogenizing) of the mean returns, the volatility of a transformed return in ISR-PCA equals the inverse SR of the respective original return. A factor analysis implemented on the ISR-PCA transformed returns identifies factors whose volatilities reflect the inverse SRs of the original returns. Accordingly, this intuition indicates that a reverse eigenvalue ranking for ISR-PCA prioritizes factors that reflect high SRs of the original returns.[1] The difference between the SR-PCA and ISR-PCA factors offers a diversification benefit to increase the maximum attainable SRs in models combining these factors.

By design, the moment-fusing constructions aim to deliver factors of significant Sharpe ratios. This focus is motivated by a rotational invariance property of the Sharpe ratios of the principal components (or, the principal factor prices). Namely, the vectors of principal factor prices in all equivalent (original and transformed) return bases are related by simple rotations (Proposition 1). When only a limited number of factors is retained (due to large data and dimensionality reduction requirement), the moment fusing prioritizes factors of high prices because the omitted factor prices (of non-retained factors) are small (as implied by the SR rotational invariance).

Our empirical analysis evaluates the in-sample (IS) and out-of-sample (OOS) pricing performance of moment-fusing constructions and related pricing factor models using FX and equity data and 4 standard pricing measures.[2] The analysis of the FX market employs the most actively traded currencies of 11 advanced economies and presents results for 4 different representative numeraire currencies (namely, the currency denominations of the US, Japan, Australia, and UK). The pricing performance of the moment-fusing models is robust to choices of numeraire curren-

[1] A reverse eigenvalue ranking prioritizes small eigenvalues. Empirically, we address the robustness issue concerning spurious small eigenvalues by applying a robust (high) lower threshold to retain only factors associated with significant eigenvalues. Conceptually, we introduce an auxiliary return basis to convert the reverse eigenvalue ranking into a standard one, which prioritizes large eigenvalues.

[2] Four pricing measures are: the maximum attainable Sharpe ratio, root-mean-square pricing error, Gibbons-Ross-Shanken (GRS) test statistic, and percentage of unexplained (idiosyncratic) return variation.

cies. The empirical analysis using FX data offers two advantages. First, the observable interest rate differentials present a known proxy for the first moment of currency returns as documented in the currency carry trade literature. Second, the change of the numeraire currency allows to assess similar moment-fusing constructions in different denomination perspectives. We find that the pricing performance of the two prominent FX benchmark factors, namely the Dollar (or, RX) and the carry trade long-short (or, HML) factors Lustig et al. (2011), varies significantly with currency denominations. The moment-fusing and related factors offer consistent pricing performance across all considered numeraire currencies. Specifically, the annualized OOS maximum attainable SRs obtained by the extended moment fusing with two retained factors are 0.28, 0.61, 0.49, and 0.61 in the numeraire currencies of USD, JPY, AUD, and GBP, and respectively are 0.09, 0.32, 0.18, and 0.27 for the benchmark model.

The analysis of the equity market employs 3 data sets, namely, the 74 and 370 extreme-decile anomaly portfolios and Fama-French 25 size and book-to-market double sorted portfolios. We use the sampled moments of these original portfolios to inform (i.e., fuse) the transformed returns. We construct moment-fusing factors on the transformed returns and also extend them with pricing restrictions (whose weights are dynamically trained). The benchmark model for the equity market employs the pricing restrictions (also dynamically trained, but without moment fusing). We retain either 3 or 5 factors in every model. For the Fama-French 25 portfolios priced by 5 retained factors, the benchmark model outperforms moment-fusing constructions, respectively delivering monthly OOS maximum attainable SRs of 0.34 and 0.28. For all other combinations of data sets and retained factor configurations, the moment-fusing or extended models deliver equal or higher OOS maximum attainable SRs than the benchmark model. This strong performance of the benchmark model shows the importance of Lettau and Pelger (2020)'s pricing restrictions for pricing factors in the equity market. It also shows that moment fusing constructions benefit from incorporating these pricing restrictions to improve the pricing performance of the combined models.

Related Literature: Our paper belongs to a large and expansive literature employing the statistical methodologies to identify systematic risk factors in the cross section of asset returns Ross (1976), in particular the covariance-based analysis (PCA) that starts with Chamberlain and Roth (1983), Connor and Korajczyk (1986), Connor and Korajczyk (1988), Litterman and Scheinkman (1991). Our paper builds on Lettau and Pelger (2020), who incorporate the mean returns into PCA by adding pricing restrictions concerning the original asset returns. While the moment fusing operates on the transformed (i.e., fused) returns, it also benefits from employing pricing restrictions to strengthen pricing performance OOS. The moment-fusing approach applies to all return data. It helps to identify pricing factors in the FX market, in which PC factors have been documented and related to prominent currency strategies in the USD denomination Lustig et al. (2011). Our paper contributes to the currency pricing literature by identifying new moment-fusing FX factors and and examining their pricing performance in various currency denominations Maurer et al. (2019). In a broader perspective, this paper is related to the literature on uncovering latent factors in empirical asset pricing using high-dimensional methods. Due to large amount of financial assets associated with various characteristics and different frictions, a zoo of factors has been identified in the literature Cochrane (2011), Harvey et al. (2016), Feng et al. (2020). Several recent approaches combine firms' or other characteristics with statistical analysis to construct proxy portfolios of enhanced exposures to (hence, elucidating) the underlying risk factors. Kelly et al. (2019) incorporate observable characteristics as instruments into the standard PCA to identify important risk factors. Kozak et al. (2018) employ Bayesian learning to enhance principal factors and identify a dimensionality reduction in a large cross-section of characteristics. Giglio and Xiu (2021) mitigate the omitted-variable bias in multi-factor risk premium estimates using PCA. Huang et al. (2022) introduce a scaled PCA that weights predictors by their predictive slopes on the target variable, thereby enhancing variables with stronger forecasting power. Giglio et al. (2025) suggest a supervised PCA method that selects test assets of significant exposures to factors of interest. Moment fusing aims to identify elusive factors by first strengthening these factors in the transformed returns. Recent aspects in assessing the performance of pricing models are discussed in Barillas and Shanken (2017) concerning the pricing error α and Kan et al. (2024) concerning the attainable Sharpe ratio. By incorporating the pricing aspects and requirements into the construction of the transformed returns, moment fusing aims to construct pricing factors that address these requirements. In the current paper, moment fusing is constructed to identify factors of high Sharpe ratios (with and without further pricing restrictions). Our empirical analysis reports both OOS α and attainable SRs (and two additional pricing measures) for all considered models.

The paper is organized as follows. Section 2 introduces the basic moment-fusing ideas and demonstrates a rotational invariance for the Sharpe ratios of PC factors. Section 3 formalizes the moment-fusing constructions and extends these constructions by incorporating pricing restrictions. Section 4 presents the empirical analysis of the moment-fusing constructions using FX data and Section 5 using equity data. Section 6 concludes. Online Appendices provide further details on empirical implementations and technical derivations.

2 Market Setup and Illustrations

This section sets the stage for the moment fusing approach applied on traded asset returns. Section 2.1 introduces a market setup of generic traded asset returns and notations. Section 2.2 presents a Sharpe ratio rotational invariance result and simple illustrations motivating the subsequent moment-fusing framework.

2.1 Market Setup and Notations

Setup: We model a generic arbitrage-free and frictionless financial market in a discrete-time setting with T periods and $T+1$ dates indexed by $t \in \{0, \ldots, T\}$, which consists of a money market account (i.e., risk-free bond) B_0 and N non-redundant generic risky assets $\{X_n\}_{n=1}^N$. In our subsequent empirical analysis of the moment-fusing framework, risky assets are currency strategies (Section 4) and equity portfolios (Section 5). For the holding period from t to $t+1$, let the bond's return and risky assets' excess returns respectively be

$$B_{0t+1} \;=\; r, \qquad X_{nt+1} \;=\; \mu_{xn} \;+\; \sigma'_{xn}\varepsilon_{xt+1}, \quad n \in \{1, \ldots, N\}, \quad t \in \{0, \ldots, T-1\}, \tag{1}$$

where r denotes the short-term risk-free rate, vector ε_{xt+1} of uncorrelated standard normal shocks represents various risks impacting asset returns, scalar μ_{xn} and vector σ_{xn} respectively denote the (excess) mean and volatilities of n-th risky asset return. In general these moments can be time-varying. For notational and exposition simplicities, we omit their time variation and time index.[3] Throughout, $\widetilde{A}$ indicates a demeaned quantity, A' a transposed quantity, $|v|$ the magnitude of a vector v, and all returns are in excess of the risk-free rate unless otherwise explicitly stated.

Transformed asset returns: Given an arbitrage-free and frictionless financial market of $N+1$ original assets $\{B_0, X_n\}_{n=1}^N$, any set of N non-redundant traded portfolios and the bond $\{B_0, Y_n\}_{n=1}^N$ constitutes an equivalent asset basis representing the same financial market. This flexibility en-

[3]The moment-fusing framework is flexible and can also incorporate (i.e., fuse) the time-varying characteristics of the original asset return moments into the construction of the transformed asset returns.

ables constructions of new asset bases that retain certain characteristics of the original one without altering the return space or the underlying model. We generically refer an equivalent asset basis (and its returns) and a transformed asset basis (and transformed asset returns).

One simple set of transformed assets concerns a pure leverage operation, in which each Y_n is a portfolio of the bond (with weight $1-\kappa_n$) and only the respective original risky asset X_n (with weight κ_n). The excess returns of these transformed assets are,[4]

$$Y_{nt+1}(\kappa_n) \;=\; \kappa_n X_{nt+1} \;=\; \underbrace{\kappa_n \mu_{xn}}_{=\mu_{yn}} + \underbrace{\kappa_n \sigma'_{xn}\varepsilon_{xt+1}}_{=\sigma'_{yn}\varepsilon_{yt+1}}, \quad n \in \{1,\ldots,N\}, \kappa_n \in \mathbb{R}. \tag{2}$$

It is important to observe that the linearly scaled excess return $\kappa_n X_{nt+1}$ (2) remains a traded excess return, has an identical SR, and perfectly correlates with the original excess return X_{nt+1}. That is, $Y_{nt+1}(\kappa_n)$ and X_{nt+1} characterize the same risk for all non-zero $\kappa_n \in \mathbb{R}$. More generally, linear combinations of traded excess returns $Y_{T\times N} = X_{T\times N} S_{N\times N}$ (where S is a full-rank matrix and X and Y contain the time series of original and transformed returns in their columns) are also traded excess returns. That is, in a frictionless financial market, return transformations can flexibly alter the covariance structure of asset returns without changing the return space or the underlying pricing model. As leading factor analysis models such as PCA center on the return covariance structure, the observations above indicate that return transformations have profound impacts on this covariance-based analysis and can help to construct important pricing factors. As an application, we employ return transformations in PCA to establish a SR rotational invariance property of PCs.

Sharpe ratio rotational invariance: To analyze the impacts of return transformations on PCA, we consider any two equivalent (original and transformed) return bases, ($\{B_0, X_n\}_{n=1}^N$ and $\{B_0, Y_n\}_{n=1}^N$) that are related by a $N\times N$ matrix A of full rank $Y_{T\times N} = X_{T\times N} S_{N\times N}$. We implement PCA on the original and transformed bases by diagonalizing respective covariance matrices, $\frac{1}{T} W'_x \widetilde{X}' \widetilde{X}\, W_x = Diag\,(\lambda_{x1},\ldots,\lambda_{xN})$ and $\frac{1}{T} W'_y \widetilde{Y}' \widetilde{Y}\, W_y = Diag\,(\lambda_{y1},\ldots,\lambda_{yN})$, where W's are $N\times N$ rotation matrices, $W'_x W_x = W_x W'_x = \mathbb{1}_{N\times N}$ and $W'_y W_y = W_y W'_y = \mathbb{1}_{N\times N}$, and $Diag\,(\{\theta_n\})$ denotes the diagonal matrix with diagonal entries $\{\theta_n\}$. Original and transformed sets of principal components (PCs) are columns of $T\times N$ matrices $\Pi_x = XW_x$ and $\Pi_y = YW_y$ respectively. Evidently,

[4] Recall that X_n denotes the excess return, so $X_n + B_0$ is the full return, of the risky asset X_n. The full return of a portfolio of the bond (with weight $1-\kappa_n$) and X_n (with weight κ_n) then is $(1-\kappa_n)B_0 + \kappa_n(X_n + B_0) = B_0 + \kappa_n X_n$. As a result, the portfolio's excess return is $\kappa_n X_n$, implying (2).

PCs are linear combinations of traded assets. Hence, PCs are also traded and constitute another two orthogonal bases ($\{B_0, \Pi_{xn}\}_{n=1}^N$ and $\{B_0, \Pi_{yn}\}_{n=1}^N$) of the return space as observed earlier.[5]

To relate original and transformed PCs and their SRs, we first employ the leverage transformation (2) to standardize these PCs, $\widehat{\Pi}_x \equiv \Pi_x Diag\left(\frac{1}{\sqrt{\lambda_x}}\right)$ and $\widehat{\Pi}_y \equiv \Pi_y Diag\left(\frac{1}{\sqrt{\lambda_y}}\right)$. As a result, N standardized and pairwise orthogonal PCs $\{\widehat{\Pi}_{xn}\}_{n=1}^N$ (which are N columns of $T \times N$ matrix $\widehat{\Pi}_x$) constitute an orthonormal basis of the risky return space, and $\{\widehat{\Pi}_{yn}\}_{n=1}^N$ similarly constitute another one.[6] Next, note that any two orthonormal bases of a N-dim vector space can be rotated from one to the other by a $N \times N$ rotation matrix $\mathcal{O}$, we have $\widehat{\Pi}_y = \widehat{\Pi}_x \mathcal{O}$, or equivalently, $\widehat{\Pi}'_y = \mathcal{O}'\widehat{\Pi}'_x$. Finally, taking the mean of every row in the last matrix equation, $\overline{\widehat{\Pi}}'_y = \mathcal{O}'\overline{\widehat{\Pi}}'_x$, and noting that the means of standardized PCs are SRs of PCs, we have,[7]

$$\underbrace{\begin{bmatrix} \frac{\overline{\Pi}_{y1}}{\sqrt{\lambda_{y1}}} \\ \vdots \\ \frac{\overline{\Pi}_{yN}}{\sqrt{\lambda_{yN}}} \end{bmatrix}}_{=SR[\Pi_y]} = \mathcal{O}' \underbrace{\begin{bmatrix} \frac{\overline{\Pi}_{x1}}{\sqrt{\lambda_{x1}}} \\ \vdots \\ \frac{\overline{\Pi}_{xN}}{\sqrt{\lambda_{xN}}} \end{bmatrix}}_{=SR[\Pi_x]}, \quad \text{or,} \quad \begin{cases} SR\left[\Pi_y\right] = \mathcal{O}' \; SR\left[\Pi_x\right], \\ \\ SR\left[\Pi_x\right] = \mathcal{O} \; SR\left[\Pi_y\right]. \end{cases} \tag{3}$$

We summarize this rotational invariance of SR in PCA in the following proposition, before discussing its implications.

Proposition 1 (Sharpe Ratio Rotational Invariance)**.** *The factor prices (i.e., SRs) of PCs constructed from any two equivalent asset bases $\{B, X_n\}_{n=1}^N$ and $\{B, Y_n\}_{n=1}^N$ of a return space are always related by a rotation, $SR\left[\Pi_y\right] = \mathcal{O}' \; SR\left[\Pi_x\right]$ (3), where $\mathcal{O}$ is a $N \times N$ rotation matrix.*

This analytical result presents important guidance for the construction of pricing factors in subsequent sections. We observe that in the difference with means and variances of principal

[5]Given that PCs $\Pi_x = XW_x$ are linear combinations of N excess returns in $X_{T\times N}$, PCs are also excess returns. Hence, the full traded return that mimics the n-th PC is $\Pi_{xn} + B_0 = (\sum_{k=1}^N X_k W_{xkn}) + B_0 = \sum_{k=1}^N (X_k + B_0) W_{xkn} + (1 - \sum_{k=1}^N W_{xkn}) B_0$, per the discussion below (2). The demeaned component of this PC-mimicking return is $\widetilde{\Pi}_{xn} = \sum_{k=1}^N \widetilde{X}_k W_{xkn}$.

[6]Specifically, N standardized PCs $\widehat{\Pi}_x \equiv \Pi_x Diag\left(\frac{1}{\sqrt{\lambda_{x1}}}, \ldots, \frac{1}{\sqrt{\lambda_{xN}}}\right)$ satisfies the orthonormal relations: $Var\left(\widehat{\Pi}_{xn}\right) = \frac{1}{T}\widetilde{\widehat{\Pi}}'_{xn}\widetilde{\widehat{\Pi}}_{xn} = 1$ and $Covar\left(\widehat{\Pi}_{xn}, \widehat{\Pi}_{xk}\right) = \frac{1}{T}\widetilde{\widehat{\Pi}}'_{xn}\widetilde{\widehat{\Pi}}_{xk} = 0$, $\forall n, k \in \{1, \ldots, N\}$ and $n \neq k$. Similar orthonormal relations hold for N standardized PCs $\widehat{\Pi}_y \equiv \Pi_y Diag\left(\frac{1}{\sqrt{\lambda_{y1}}}, \ldots, \frac{1}{\sqrt{\lambda_{yN}}}\right)$.

[7]To compute the mean of every row of the matrix $\widehat{\Pi}'_y = \mathcal{O}'\widehat{\Pi}'_x$, we multiply to the left of both sides by $\frac{1}{T}\mathbb{1}_{T\times 1}$, obtaining $\overline{\widehat{\Pi}}'_y = \mathcal{O}'\overline{\widehat{\Pi}}'_x$, where the $N \times 1$ column $\overline{\widehat{\Pi}}'_x$ contains means of N standardized PCs $\{\widehat{\Pi}_{xn}\}_{n=1}^N$, and $\overline{\widehat{\Pi}}'_y$ contains means of N standardized PCs $\{\widehat{\Pi}_{yn}\}_{n=1}^N$. Since these standardized PCs are constructed as $\widehat{\Pi}_x = \Pi_x Diag\left(\frac{1}{\sqrt{\lambda_x}}\right)$, their means are $\overline{\widehat{\Pi}}_{xn} = \frac{\overline{\Pi}_{xn}}{\sqrt{\lambda_{xn}}}$ which are the SRs of PCs Π_{xn}, $n \in \{1, \ldots, N\}$. Similarly, the means of standardized PCs $\widehat{\Pi}_{yn}$ are the SRs of PCs Π_{yn}, $n \in \{1, \ldots, N\}$. These results deliver (3).

factors, which can be linearly scaled by arbitrary factors (2), principal factors' prices (or SRs of any set of principal components of a return space) are subject to a rotational variance. As a result, when only a limited number of principal factors is retained (i.e., dimensionality reduction), this rotational invariance assures that, for any asset return basis, the higher prices the retained factors have, the lower prices the omitted (not retained) factors have. That is, Proposition 1 indicates that a combination of return transformations and PCA enables the dimensionality reduction by identifying a sparse structure of factor prices. In such a structure, retaining few factors of dominant risk prices offers high attainable SRs and robustness against scaling operations that can upset the covariance-based ranking and choice of factors.

2.2 Moment Fusing and Preliminary Illustrations

Moment fusing is a general, systematic and flexible framework to incorporate information from important moments (e.g., mean, variance, skewness) of asset returns into the construction of pricing factors. It consists of two stages. In the first stage, we construct transformed asset returns $Y_{T\times N}$ by incorporating the first (and higher) moments m_x of original returns $X_{T\times N}$. As a result, the covariance structure $\Sigma_y(m_x) = \frac{1}{T}\widetilde{Y}'(m_x)\widetilde{Y}(m_x)$ of the transformed returns is informed by important moments of original returns. In the second stage, a covariance-based factor analysis (e.g., PCA) is performed on the transformed returns to identify factors $\Pi_y(m_x)$ that reflect original moments and price original returns.

Moment Fusing:

$$
\begin{aligned}
&\text{1st stage – Transformation} \quad &&\begin{cases} X \longrightarrow Y(m_x) = XS(m_x); \\ \Sigma_x \longrightarrow \Sigma_y(m_x) = S'(m_x)\Sigma_x S(m_x). \end{cases} \\
&\text{2nd stage – Factor analysis} \quad &&\begin{cases} W_y'(m_x)\left[\Sigma_y(m_x)\right]W_y(m_x) = Diag\left[\lambda_y(m_x)\right] \\ \Longrightarrow \Pi_y(m_x) = Y(m_x)W_y(m_x). \end{cases}
\end{aligned}
\tag{4}
$$

The $N \times N$ full-rank moment-fusing matrix $S(m_x)$ is critical for the construction. It is fused with, and hence is a function of, moments m_x of the original returns. Compared to the factor analysis implemented on the original returns (i.e., diagonalizing Σ_x), the one implemented on the transformed returns (i.e., diagonalizing $\Sigma_y(m_x) = S'(m_x)\Sigma_x S(m_x)$) also concerns moments m_x other than the second moment Σ_x. It therefore also results in informed pricing factors $\Pi_y(m_x)$. In the moment-fusing framework, the first stage preserves the return space (hence, the underlying model), the second stage retains and utilizes advantages of the powerful factor analysis methodology. Below, we present three simple and known asset transformations and discuss why they do not yet have desirable moment-fusing properties. The discussion serves as preliminary illustrations and motivation for elaborate moment-fusing constructions in the next section.

PCA using correlation matrix: The correlation matrix of original returns can be constructed as the covariance matrix of transformed returns, which are scaled to have an unit volatility. Employing $\kappa_n = \frac{1}{|\sigma_{xn}|}$ in the transformation (2) produces

$$Y_{nt+1}(\kappa_n) = \kappa_n X_{nt+1} = \frac{\mu_{xn}}{|\sigma_{xn}|} + \underbrace{\frac{\sigma'_{xn}}{|\sigma_{xn}|}}_{=\widehat{\sigma}'_{xn}} \varepsilon_{xt+1}, \qquad n \in \{1, \ldots, N\}, \tag{5}$$

where $\widehat{\sigma}_{xn}$ is an unit vector ($|\widehat{\sigma}_{xn}| = 1$). Evidently, $Corr_t\left[Y_{nt+1}(\kappa_n), Y_{kt+1}(\kappa_k)\right] = Cov_t\left[X_{nt+1}(\kappa_n), X_{kt+1}(\kappa_k)\right]$, for all $n, k \in \{1, \ldots, N\}$. Therefore, a PCA implemented on the correlation matrix of the original asset returns is identical to a standard PCA implemented on the covariance matrix of the transformed asset returns. Note that the operation (5) does not constitute a moment-fusing transformation because the scaling parameters $\kappa_n = \frac{1}{|\sigma_{xn}|}$, $n \in \{1, \ldots, N\}$, do not incorporate the first moment of original asset returns. While the PCA implemented on the transformed asset returns identifies principal factors that explain the covariation of $\{Y_{nt+1}(\kappa_n)\}$, these factors are not informed by the pricing of the original assets. As a result, leading principal factors obtained from a correlation matrix analysis are not necessarily the important pricing factors in the given return space.

PCA using uncentered covariance matrix: The standard (centered) covariance matrix $\widetilde{X}'\widetilde{X}$ quantifies the return comovements around the means of respective asset returns and employs exclusively the demeaned asset returns $\{\widetilde{X}_n\}_{n=1}^N$. By construction, the centered covariance matrix and a standard PCA analysis based on this matrix are stripped of information about the first moment of asset returns. In contrast, the uncentered covariance matrix $X'X$ employs non-demeaned asset returns $\{X_n\}_{n=1}^N$ and hence contains information about the first moment of asset returns. Hence, the employment of non-demeaned returns in defining the uncentered covariance matrix constitutes a simple fusing of the first moment of asset returns into the factor analysis based on this matrix.

However, such a simple moment fusing does not necessary identify important pricing factors because while containing information about the first moment, uncentered covariances $X'X$ do not delineate risks (i.e., return volatilities) from risk premia (i.e., mean excess returns). The mixing of these two quantities complicates an understanding of the tradeoff between risks and compensated returns in an asset pricing analysis. Intuitively, note that the uncentered covariance matrix $X'X$ quantifies the return comovements around zero value of returns, whereas the centered one $\widetilde{X}'\widetilde{X}$ quantifies the comovements of returns with respect to the means of these returns. In case the means of returns are similar, $X'X$ and $\widetilde{X}'\widetilde{X}$ capture similar risks. But when the means of returns are sufficiently heterogeneous, comovements characterized by the uncentered $\widetilde{X}'\widetilde{X}$ do not not simply or purely reflect common shocks inherent in these returns (but also the similarities in the cross section of returns' means). As a result, an important PC (associated with a larger eigenvalue of $X'X$) may capture a spurious common pattern in the cross section of returns' means. Such an ambiguous fusing of returns' first moments into the uncentered covariance matrix does not necessarily identify important pricing factors.

Rotation transformation: Finally, we consider the transformed returns which are obtained the original ones by a rotation, $Y_{T\times N} = X_{T\times N}Q_{N\times N}$, where $Q_{N\times N}$ is a rotation matrix. In this case, original and transformed returns have identical PCs, $\Pi_x = \Pi_y$, and hence, identical principal risk factors.[8] That is, even when we fuse the rotation matrix Q with information about the first moment of original asset returns, such moment fusing has no effects on principal factors obtained from the transformed returns. Intuitively, this is because when every original return in X is rotated rigidly by the same matrix Q, its covariance structure remain the same, resulting in the same PCs for the two (original and transformed) return bases X and Y. As as result, a rotation transformation is not a moment-fusion construction due to the rotation rigidity that keeps intact relationships between the first and second moments of asset returns.

These simple illustrations of the PCA implemented on the correlation and uncentered covariance matrices and the rotation transformation motivate more sophisticated fusing constructions that do not simply incorporate the first moment of original returns into the asset transformation but also do that in a way to inform about the underlying asset pricing model. We presents such moment-fusing constructions next.

[8]The covariance matrices of transformed and original returns are related as, $\Sigma_y = \frac{1}{T}\widetilde{Y}'\widetilde{Y} = \frac{1}{T}Q'\widetilde{X}'\widetilde{X}Q = Q'\Sigma_x Q$. If Σ_x is diagonalized by W_x, or $W_x'\Sigma_x W_x = Diag\,(\lambda_x)$, then Σ_y is diagonalized by $W_y = Q'W_x$. Indeed, $(W_y')(\Sigma_y)(W_y) = (W_x'Q)(Q'\Sigma_x Q)(Q'W_x) = W_x'\Sigma_x W_x = Diag\,(\lambda_x)$. PCs associated with the transformed returns are then identical to those associated with the original returns, $\Pi_y = YW_y = (XQ)(Q'W_x) = XW_x = \Pi_x$.

3 Moment Fusing: Main Constructions

This section introduces the main moment-fusing constructions of the paper by transforming the return volatilities into SRs (Section 3.1), standardizing the mean returns (Section 3.2), and relating and combining these constructions with the pricing restrictions of the literature (Section 3.3).

3.1 Sharpe Ratio Matrix and SR-PCA

We consider a moment-fusing construction that aims to identify factors of significant prices equalizes the second moment of the transformed asset returns directly with Sharpe ratio of the original asset returns. Specifically, the first stage of the construction employs scaling parameters $\kappa_n = \frac{(\mu_{xn})^{\frac{1}{2}}}{|\sigma_{xn}|^{\frac{3}{2}}}$ in (2) and produces the transformed returns

$$Y_{nt+1}(\kappa_n) = \kappa_n X_{nt+1} = \underbrace{(SR\,[X_n])^{\frac{3}{2}}}_{=\mu_{yn}} + \underbrace{(SR\,[X_n])^{\frac{1}{2}}}_{=\sigma_{yn}} \widehat{\sigma}'_{xn}\varepsilon_{xt+1}, \tag{6}$$

where $SR\,[X_n] = \frac{\mu_{xn}}{|\sigma_{xn}|}$ is the Sharpe ratio of nth original return. In this construction, transformed assets' return variances equal the corresponding original returns' Sharpe ratios, while their mean returns equal the power $\frac{3}{2}$ of the corresponding original returns' Sharpe ratios. In the general notation (4), the moment-fusing matrix is $S(m_x) = Diag\left(\frac{(\mu_x)^{\frac{1}{2}}}{|\sigma_x|^{\frac{3}{2}}}\right)$, and the transformed covariance matrix Σ_y is related to the original counterpart Σ_x as

$$\frac{1}{T}\widetilde{Y}'\widetilde{Y} = \begin{bmatrix} \frac{(\mu_{x1})^{\frac{1}{2}}}{|\sigma_{x1}|^{\frac{3}{2}}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{(\mu_{xN})^{\frac{1}{2}}}{|\sigma_{xN}|^{\frac{3}{2}}} \end{bmatrix} \frac{\widetilde{X}'\widetilde{X}}{T} \begin{bmatrix} \frac{(\mu_{x1})^{\frac{1}{2}}}{|\sigma_{x1}|^{\frac{3}{2}}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{(\mu_{xN})^{\frac{1}{2}}}{|\sigma_{xN}|^{\frac{3}{2}}} \end{bmatrix},$$

$$\text{or,} \qquad \Sigma_y = Diag\left[\frac{(\mu_{xn})^{\frac{1}{2}}}{|\sigma_{xn}|^{\frac{3}{2}}}\right] \Sigma_x \; Diag\left[\frac{(\mu_{xN})^{\frac{1}{2}}}{|\sigma_{xn}|^{\frac{3}{2}}}\right] \equiv SRM(X). \tag{7}$$

Hereafter, we refer to $\Sigma_y = SRM(X)$ defined above as the Sharpe ratio matrix of original returns $\{X_n\}_{n=1}^N$. It generalizes the Sharpe ratio $\frac{\mu_n}{|\sigma_n|}$ for a single asset return to a matrix version for N asset returns. Evidently, $SRM(X)$ (7) is a moment-fusing construction as the original means $\{\mu_{xn}\}_{n=1}^N$ are incorporated into the transformed volatilities $\{\sigma_{yn}\}_{n=1}^N$ (6) in such a way that $\sigma_{yn} =$

$\frac{(\mu_{xN})^{\frac{1}{2}}}{|\sigma_{xn}|^{\frac{1}{2}}} = (SR[X_n])^{\frac{1}{2}}$. Eigenvalues and eigenvectors of Σ_y then are also fused with original mean returns.

Next, the second stage of the construction implements the PCA on the transformed returns (diagonalizing Σ_y) which is equivalent to a factor analysis based on the Sharpe ratio matrix of the original returns (diagonalizing $SRM(X)$). Intuitively, factors prioritized by their eigenvalues are dominated by large common components of transformed return volatilities, which are also large components of original Sharpe ratios (6). Therefore, a standard eigenvalue ranking concerning $\Sigma_y = SRM(X)$ tends to prioritize factors of significant Sharpe ratios (or factor prices) for the original asset returns. Building on this intuition and applying PCA on the SR matrix $SRM(X)$ (7), the moment-fusing construction prioritizes more volatile factors among PCs $\{\Pi_{yn}\}$ associated with the transformed return basis $\{Y_n\}$ (6),

$$\textbf{Sharpe Ratio-PCA (SR-PCA):} \qquad \lambda_{yn} > \lambda_{yk} \longrightarrow \Pi_{yn} \succ \Pi_{yk}, \tag{8}$$

where $\{\lambda_{yn}\}$ are eigenvalues of the inverse SR matrix $SRM(X)$ and $\{\Pi_{yn}\}$ are the associated PCs. We refer to this moment-fusing construction as the Sharpe Ratio-PCA (SR-PCA). Two important observations concerning the moment-fusing construction SR-PCA are in order.

First, as the transformed mean returns $\mu_{yn} = (SR\,[X_n])^{\frac{3}{2}}$ (6) increase with Sharpe ratios of the corresponding original returns, the SR-PCA moment fusing aligns the contributions of the first and second moments of risky assets to the factors. That is, all else being equal, as more volatile transformed returns (larger $\sigma_{yn} = (SR\,[X_n])^{\frac{1}{2}}$) have stronger influences on the prioritized SR-PCA factors, their higher risk premia (larger $\mu_{yn} = (SR\,[X_n])^{\frac{3}{2}}$) also enhance the prices of these factors. Such an alignment prevents the first moment to adversely skew the factors and their factor prices, allowing SR-PCA to identify principal factors of significant factor prices.

Second, given the transformed return volatilities $\sigma_{yn} = (SR\,[X_n])^{1/2}$, $\forall n \in \{1, \ldots, N\}$, (6), the total variance of all transformed returns equal the total SR of all original assets, $\sum_{n=1}^{N} \sigma_{yn}^2 = \sum_{n=1}^{N} SR\,[X_n]$. Note that this total variance also equal the sum of all eigenvalues associated with the transformed returns,[9] or $\sum_{n=1}^{N} \lambda_{yn} = \sum_{n=1}^{N} \sigma_{yn}^2 = \sum_{n=1}^{N} SR\,[X_n]$. Hence, when we prioritize and retain $K < N$ important PCs $\{\Pi_{yi}\}_{i=1}^{K}$ in accordance with the ranking (8), the percentage of

[9]This is a standard PCA property. As returns $\{Y_n\}_{n=1}^{N}$ and PCs $\{\Pi_y\}_{n=1}^{N}$ are related by a rotation matrix, $\Pi_y = YW_y$ or $Y = \Pi_y W_y'$, their total variances are the same, $\sum_{n=1}^{N} \sigma_{yn}^2 = Trace\{\widetilde{Y}'\widetilde{Y}\} = Trace\{W_y \widetilde{\Pi}_y' \widetilde{\Pi}_y W_y'\} = Trace\{\widetilde{\Pi}_y' \widetilde{\Pi}_y W_y' W_y\} = Trace\{\widetilde{\Pi}_y' \widetilde{\Pi}_y\} = Trace\{Diag \lambda_y\} = \sum_{n=1}^{N} \lambda_{yn}$.

the total return covariation that is explained by these K PCs is,[10]

$$\frac{\sum_{n=1}^{N} Cov\left(Y_n\,,\,\sum_{i=1}^{K}\Pi_{yi}W'_{yin}\right)}{\sum_{n=1}^{N} Var\left(Y_n\right)} = \frac{\sum_{i=1}^{K}\lambda_{yi}}{\sum_{n=1}^{N}\lambda_{yn}} = \frac{\sum_{i=1}^{K}\lambda_{yi}}{\sum_{n=1}^{N} SR\left[X_n\right]}. \tag{9}$$

In the sense of the ratio (9), the SR-PCA ranking (8) assures that a SR-PCA factors Π_{yi} associated with a larger value λ_{yi} characterizes a more prominent common component of the transformed returns' volatilities (equivalently, the original returns' SRs), quantified as a larger explained portion of the total original assets' SRs. Combining this characterization with the SR rotational invariance (Proposition 1), which aligns the retention of PCs with high factor prices with the reduction of the omitted (non-retained) SRs, presents the rationale for the SR-PCA ranking (8). Namely, this ranking aims to identify and prioritize SR-PCA factors that optimize over the SRs of the original assets via maximizing the explained covariation of the transformed returns by these factors.

3.2 Inverse Sharpe Ratio Matrix and ISR-PCA

Another moment-fusing construction starts with the observation that as the covariance-based factor analysis ignores the information about the mean returns, the transformation from the original returns to principal factors may have unintended influences on factors' risk premia. For instance, volatile principal factors may load little on some assets of high mean returns because these loadings are determined exclusively from how these assets contribute to the common movements of all asset returns. To address these unintended influences, we consider a moment-fusing construction in which we standardize all mean returns to a same notional value μ in the first stage, and implement the factor analysis on the transformed (standardized) returns. The construction aims to control for (i.e., homogenize) return first moment, transforming and organizing differences among returns into their second moment, which then can be appropriately analyzed by the covariance-based PCA approach.

Specifically, the first stage of the construction employs scaling parameters $\kappa_n = \frac{\mu}{\mu_{xn}}$ in (2)

[10]Note that in the PCA model of the second stage, asset returns and PCs are related by $Y = \Pi_y W'_y$. Therefore, $\sum_{i=1}^{K}\Pi_{yi}W'_{yin}$ is the model-implied part of the return Y_n. The first equality in (9) is a standard result for the percentage of the explained return covariation in PCA. The second equality arises from SR-PCA transformation property just derived above $\sum_{n=1}^{N}\lambda_{yn} = \sum_{n=1}^{N} SR\left[X_n\right]$. See also the percentage of unexplained variation in returns (33).

and produces the transformed returns (below, $SR\left[X_n\right]=\frac{\mu_{xn}}{|\sigma_{xn}|}$, $\forall n$, as in (6))

$$Y_{nt+1}(\kappa_n)=\kappa_n X_{nt+1}=\mu+\mu\,\frac{1}{SR\left[X_n\right]}\widehat{\sigma}'_{xn}\varepsilon_{xt+1}, \tag{10}$$

$$\Longrightarrow \quad \mu_{yn}=\mu, \quad \sigma_{yn}=\mu\frac{|\sigma_{xn}|}{\mu_{xn}}=\mu\frac{1}{SR\left[X_n\right]}, \quad n\in\{1,\ldots,N\},$$

where μ is a notional constant parameter.[11] In this construction, up to the non-material multiplicative constant μ, the transformed return volatilities equal the inverse of the corresponding original Sharpe ratios, $\sigma_{yn}\sim\frac{1}{SR[X_n]}$, $\forall n\in\{1,\ldots,N\}$. In the general notation (4), the moment-fusing matrix is $S(m_x)=Diag\left(\frac{1}{\mu_x}\right)$, and the transformed covariance matrix Σ_y is related to the original counterpart Σ_x as (up to μ^2)

$$\frac{1}{T}\widetilde{Y}'\widetilde{Y}=\begin{bmatrix}\frac{1}{\mu_{x1}} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \frac{1}{\mu_{xN}}\end{bmatrix}\frac{\widetilde{X}'\widetilde{X}}{T}\begin{bmatrix}\frac{1}{\mu_{x1}} & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \frac{1}{\mu_{xN}}\end{bmatrix}, \quad \text{or,} \quad \Sigma_y=\underbrace{Diag\left[\frac{1}{\mu_x}\right]\ \Sigma_x\ Diag\left[\frac{1}{\mu_x}\right]}_{\equiv ISRM(X)}, \tag{11}$$

Hereafter, we refer to $\Sigma_y=ISRM(X)\equiv Diag\left[\frac{1}{\mu_x}\right]\ \Sigma_x\ Diag\left[\frac{1}{\mu_x}\right]$ as the inverse Sharpe ratio matrix of original returns $\{X_n\}_{n=1}^N$. It generalizes the inverse of (squared) Sharpe ratio $\frac{|\sigma|^2}{\mu^2}$ for a single asset return to a matrix version for N asset returns. Evidently, $ISRM(X)$ (11) is a moment-fusing construction as the original means $\{\mu_{xn}\}_{n=1}^N$ are incorporated into the transformed volatilities $\{\sigma_{yn}\}_{n=1}^N$ (10). Eigenvalues and eigenvectors of Σ_y then are also fused with original mean returns. Note that the moment-fusing pattern of $ISRM(X)$ (11) differs from the previous one of $SRM(X)$ (7). This is evidenced in the fact that while being related, $ISRM(X)$ strictly differs from $[SRM(X)]^{-2}$.

Next, the second stage of the construction implements the PCA on the transformed returns (diagonalizing Σ_y) which is equivalent to a factor analysis based on the inverse Sharpe ratio matrix of the original returns (diagonalizing $ISRM(X)$). Intuitively, factors prioritized by their eigenvalues are dominated by large common components of transformed return volatilities, which are small components of original Sharpe ratios (10), and vice versa. This intuition indicates that a reverse eigenvalue ranking concerning $\Sigma_y=ISRM(X)$ then tends to prioritize factors of significant Sharpe ratios (or factor prices) for the original asset returns (hence, the

[11] The specific value of μ has no material effect on the moment-fusing procedure or the resulting factors as it is a common multiplicative factor for all returns $\{Y_n\}_1^N$.

name inverse SR matrix). Building on this intuition and applying PCA on the inverse SR matrix $ISRM(X)$ (11), the moment-fusing construction prioritizes less volatile factors among PCs $\{\Pi_{yn}\}$ associated with the transformed return basis $\{Y_n\}$ (10),

$$\textbf{Inverse Sharpe Ratio-PCA (ISR-PCA):} \qquad \lambda_{yn} < \lambda_{yk} \longrightarrow \Pi_{yn} \succ \Pi_{yk}, \tag{12}$$

where $\{\lambda_{yn}\}$ are eigenvalues of the inverse SR matrix $ISRM(X)$ and $\{\Pi_{yn}\}$ are the associated PCs. We refer to this moment-fusing construction as the inverse Sharpe Ratio-PCA (ISR-PCA). For the robustness against possible spurious factors associated with small eigenvalues of $ISRM(X)$, our empirical analysis (Sections 4 and 5) retains only eigenvalues $\{\lambda_{ny}\}$ above a threshold sufficiently different from zero, before applying the ISR-PCA prioritization (12). Alternatively, due to the flexibility of the moment-fusing construction, the reverse eigenvalue ranking and ISR-PCA can be amended by introducing N auxiliary returns, $\widetilde{Z} \equiv \widetilde{X}\,[\Sigma_x]^{-1}\,Diag[\mu_x]$. The auxiliary covariance matrix equals the inverse of $ISRM(X)$, or $\Sigma_z = \Sigma_y^{-1} = [ISRM(X)]^{-1}$. The reverse eigenvalue ranking (prioritizing small eigenvalues) of $ISRM(X)$ can be implemented by a standard eigenvalue ranking (prioritizing large eigenvalues) of Σ_z, obtaining ISR-PCA factors $\Pi_y = YW_y$ in the second stage of the moment-fusing construction.[12] Finally, we recall that the transformed mean returns (10) are identical. Intuitively, the ISR-PCA moment fusing homogenizes the contribution of the first moment of all risky assets $\{Y_n\}_{n=1}^N$, leaving their second moments (i.e., the inverse SRs of original returns) and the associated PCA to identify principal factors of significant factor prices.

3.3 Pricing Restrictions and Moment Fusing

The moment fusing constructions considered above are flexible to incorporate some important features of other pricing approaches. One such prominent approach is Lettau and Pelger (2020)'s Risk-premium principal component analysis (RP-PCA). The RP-PCA takes into account the first moment of asset returns in the form of pricing restrictions, which are the penalties imposed on the deviation between the empirical mean excess returns and the risk premia implied by RP-PCA

[12] Specifically, because matrix Σ_z and its inverse $\Sigma_z^{-1} = \Sigma_y$ are diagonalized by the same rotation matrix W_y, we have $Diag\,(\lambda_z) = W_y'\Sigma_z W_y = W_y'\Sigma_y^{-1} W_y = Diag\left(\frac{1}{\Lambda_y}\right)$. After this diagonalization of Σ_z (and adopting the standard eigenvalue ranking for $\{\lambda_{zn}\}$, which is equivalent to the reverse eigenvalue ranking for $\{\lambda_{yn}\}$), we obtain the rotation matrix W_y and construct the transformed assets as $\Pi_y = YW_y$, wherein W_y contains the fused information about the first moments (and SRs) of the original return X.

factors. In this section, we first briefly discuss the relation between RP-PCA and moment-fusing frameworks. We then combine the two approaches to incorporate the pricing restrictions into ISRM-PCA and SRM-PCA constructions.

RP-PCA and moment fusing: Given asset returns $\{X_{nt}\}_{n=1}^{N}$ (1), RP-PCA searches for pricing factors $\{\Pi_{nt}^{RP}\}_{n=1}^{N}$ by minimizing an weighted objective function of unexplained covariation (concerning the second moment) and pricing errors (i.e., pricing restrictions, concerning the first moment) of asset returns

$$\min_{\{\Pi^{RP}, W^{RP}\}} \text{Trace} \left\{ \frac{1}{T} \left[\widetilde{X} - \widetilde{\Pi}^{RP} W^{RP\prime} \right]' \left[\widetilde{X} - \widetilde{\Pi}^{RP} W^{RP\prime} \right] \; + \; (1+\gamma) \left[\overline{X} - \overline{\Pi}^{RP} W^{RP\prime} \right]' \left[\overline{X} - \overline{\Pi}^{RP} W^{RP\prime} \right] \right\}, \tag{13}$$

where $\widetilde{X}_{T\times N}$ and $\overline{X}_{1\times N}$ denote respectively matrices of (demeaned) asset returns and their means, $\widetilde{\Pi}^{RP}_{T\times N}$, $\overline{\Pi}^{RP}_{1\times N}$, and $W^{RP}_{N\times N}$ the matrices of (demeaned) factors, their means, and loadings. Parameter γ specifies the relative weight between pricing errors and unexplained covariation in the objective function.[13] Lettau and Pelger (2020) show that RP-PCA factor construction is equivalent to diagonalizing a $N \times N$ covariance matrix $\Sigma_x^{RP}(\gamma)$ adjusted by the pricing restrictions. Accordingly, the prioritization of RP-PCA factors also reduces to the standard eigenvalue ranking concerning the adjusted covariance matrix $\Sigma_x^{RP}(\gamma)$ (Lettau and Pelger (2020))

$$\Pi^{RP}(\gamma) = X W^{RP}(\gamma), \quad \text{with} \quad \begin{cases} \Sigma_x^{RP}(\gamma) \equiv \frac{1}{T} \widetilde{X}' \widetilde{X} + (1+\gamma)\, \overline{X}'\, \overline{X}, \\ W^{RP\prime}(\gamma)\, \Sigma_x^{RP}(\gamma)\, W^{RP}(\gamma) = Diag\left[\lambda_n^{RP}(\gamma)\right], \end{cases} \tag{14}$$

$$\textbf{Risk-Premium PCA (RP-PCA):} \quad \lambda_n^{RP}(\gamma) > \lambda_k^{RP}(\gamma) \longrightarrow \Pi_n^{RP}(\gamma) \succ \Pi_k^{RP}(\gamma).$$

Intuitively, a deviation between mean original returns and their RP-PCA implied risk premia (i.e., implied by their loadings on RP-PCA factors) are transformed into a reduced covariation quantifies by the adjusted covariance matrix $\Sigma_x^{RP}(\gamma)$. As a result, volatile PCs associated with $\Sigma_x^{RP}(\gamma)$ also characterizes pricing factors that generate reduced price errors for original asset returns. Hereafter, we omit the explicit γ-dependence notation attached to RP-PCA's quantities when the omission does not create ambiguity.

While both RP-PCA and moment-fusing frameworks incorporate information about the first moment of asset returns, their factors differ. At the construction level, RP-PCA factors are con-

[13] In comparison, the standard PCA identifies factors (or principal components) by minimizing only the unexplained covariation of asset returns, $\min_{\{\Pi^{RP}, W^{RP}\}} \text{Trace} \left(\left[\widetilde{X} - \widetilde{\Pi}^{RP} W^{RP\prime} \right]' \left[\widetilde{X} - \widetilde{\Pi}^{RP} W^{RP\prime} \right] \right)$. As a result, RP-PCA coincide with PCA when $\gamma = -1$.

structed directly from the original asset returns $\{B_0, X_n\}_{n=1}^N$ by effectively adjusting their covariance structure Σ_x^{RP} (14), resulting in correlated factors $\{\Pi_n^{RP}\}_{n=1}^N$.[14] In contrast, moment-fusing factors are constructed indirectly from the original returns by first constructing the transformed returns via the moment-fusing matrix $S(m_x)$ (4), resulting in uncorrelated factors as PCs $\{\Pi_{yn}(m_x)\}_{n=1}^N$ of the transformed returns. More importantly, the subjective functions underlying RP-PCA and moment-fusing factors are different. RP-PCA factors (14) minimize a linear combination of first (pricing errors) and second (unexplained covariation) moments. Moment-fusing ISR-PCA (12) and SR-PCA (8) prioritize non-linear combinations (in the form of inverse SRs and SRs) of the two moments. Whether RP-PCA factors prioritize SRs can be explicitly seen by computing the prices of these factors,[15]

$$SR\left(\Pi_n^{RP}\right) = \frac{\mu_{\Pi n}}{\sqrt{Var\left(\Pi_n^{RP}\right)}} = \left[\frac{\lambda_n^{RP}}{\mu_{\Pi n}^2} - (\gamma+1)\right]^{-\frac{1}{2}}, \tag{15}$$

where $\mu_{\Pi n}$ is the mean return of the n-th RP-PCA factor, $\mu_{\Pi n} = \overline{X} W_n^{RP}$ and W_n^{RP} is the n-th column of rotation matrix W^{RP} (14). Evidently, RP-PCA factor price $SR\left(\Pi_n^{RP}\right)$ is inversely related to the ratio $\frac{\lambda_n^{RP}}{\mu_{\Pi n}^2}$, but not to λ_n^{RP} alone. As a result, RP-PCA factors, which are prioritized exclusively on the magnitude of λ_n^{RP}, do not necessarily have largest prices $SR\left(\Pi_n^{RP}\right)$.[16] The non-monotonic relationship between RP-PCA SRs and eigenvalues indicates improvements by combining RP-PCA and moment-fusing approaches.

Combining pricing restrictions with moment fusing: We extend the moment-fusing framework to incorporate Lettau and Pelger (2020)'s pricing restriction into the second (factor analysis) stage of the framework (4). Specifically, after constructing the transformed returns in the first stage, we optimize the weighted objective function (13) using the transformed returns $Y = XS(m_x)$. That is, we replace the original X by $XS(m_x)$ in that objective function. Next, by em-

[14]Note that the covariance matrix of RP-PCA factors $\frac{1}{T}\widetilde{\Pi}^{RP\prime}\widetilde{\Pi}^{RP} = W^{RP\prime}\frac{1}{T}\widetilde{X}'\widetilde{X}W^{RP} = W^{RP\prime}\Sigma_x W^{RP}$ is not a diagonal matrix because the orthogonal matrix diagonalizes the adjusted covariance matrix Σ_x^{RP}, but not the original Σ_x.

[15]Recall that λ_n^{RP} is not the variance of n-th RP-PCA factor Π_n^{RP}. Instead, $Var\left(\Pi_n^{RP}\right) = E\left[\left(\Pi_n^{RP}\right)^2\right] - E^2\left[\Pi_n^{RP}\right] = \frac{1}{T}\Pi_n^{RP\prime}\Pi_n^{RP} - \mu_{\Pi n}^2 = W_n^{RP\prime}\frac{1}{T}X'XW_n^{RP} - \mu_{\Pi n}^2$. To compute this variance, note from (14) that $X'X = \frac{1}{T}\widetilde{X}'\widetilde{X} + \overline{X}'\overline{X} = \Sigma_x^{RP} - \gamma\overline{X}'\overline{X}$, which then implies $W_n^{RP\prime}\frac{1}{T}X'XW_n^{RP} = \lambda_n^{RP} - \gamma\mu_{\Pi n}^2$. As a result, the variance of RP-PCA factor is, $Var\left(\Pi_n^{RP}\right) = \lambda_n^{RP} - (\gamma+1)\mu_{\Pi n}^2$, which delivers the factor price in (15).

[16]Note that the RP-PCA minimization of pricing errors and unexplained covariation in asset returns (13) is equivalently characterized by the diagonalization (14), which is not equivalent to the determination of factor risk premia $\{\mu_{\Pi n}\}$. That is, leading RP-PCA factors (associated with larger λ_n^{RP}) do not necessarily have highest factor risk premia $\mu_{\Pi n}$. As a result, when $\mu_{\Pi n}$ do not increase adequately with $\sqrt{\lambda_n^{RP}}$, factor prices $SR\left(\Pi_n^{RP}\right)$ do not increase with λ_n^{RP}. In this premise, it is possible that RP-PCA factor ranking does not align with the ranking based on factor prices. The choice of weight γ and the combination of RP-PCA and moment fusing help mitigate this issue.

ploying a similar deduction from the optimization (13) to the diagonalization of an adjusted covariance matrix (14), the second stage of the moment-fusing framework (4) extended with pricing restrictions amounts to diagonalizing the adjusted matrix $\Sigma_y(m_x) = \frac{1}{T}\widetilde{Y}'\widetilde{Y} + (1+\gamma)\,\overline{Y}'\,\overline{Y}$.

In summary, combining the moment fusing (4) with RP-PCA (14), in terms of original returns $X_{T\times N}$, the pricing factors in the $T \times N$ matrix $\Pi_y(\gamma, S_m)$ are constructed as follows

$$\Pi_y(\gamma, m_x) = XS_m\, W_y(\gamma, m_x), \qquad \begin{cases} \Sigma_y(\gamma, m_x) \equiv S'(m_x)\left[\frac{1}{T}\widetilde{X}'\widetilde{X} + (1+\gamma)\,\overline{X}'\,\overline{X}\right] S(m_x), \\ {W_y}'(\gamma, m_x)\,\Sigma_y(\gamma, m_x)\,W_y(\gamma, m_x) = Diag\left[\lambda_{yn}(\gamma, m_x)\right], \end{cases}$$

$$\textbf{Extended moment fusing:}\quad \lambda_{yn}(\gamma, m_x) > \lambda_{yk}(\gamma, m_x) \longrightarrow \Pi_{yn}(\gamma, m_x) \succ \Pi_{yk}(\gamma, m_x). \tag{16}$$

Notationally, the constructed factors $\Pi_y(\gamma, m_x)$ feature both moment fusing (m_x) and pricing restrictions (γ) attributes. Intuitively, the pricing restrictions discipline pricing errors that persist among transformed returns $Y(m_x)$ that are fused with original moments m_x. As an implication of Proposition 1, factors of significant factor prices (or, SRs) obtained in this procedure (concerning the set of transformed returns Y) also deliver significant maximum attainable Sharpe ratio in the set of original returns X because these two sets are equivalent bases. Empirically, we test the constructed pricing factors $\Pi_y(\gamma, m_x)$ against original returns X. We further employ a training procedure to determine γ dynamically to obtain the optimal out-of-sample attainable Sharpe ratios (detailed in Appendix A.1).

When applying the above extended moment fusing construction to the SR-PCA (7), we employ the respective moment-fusing matrix $S(m_x) = Diag\left(\frac{(\mu_x)^{\frac{1}{2}}}{|\sigma_x|^{\frac{3}{2}}}\right)$ in the construction (16). When applying the extended moment fusing construction to the ISR-PCA (11), we employ the respective $S(m_x) = Diag\left(\frac{1}{\mu_x}\right)$, adopting a negative sign associated with the covariation term $\widetilde{X}'\widetilde{X}$ in the adjusted covariance $\Sigma_y(\gamma, m_x)$ in (16) and using the standard eigenvalue ranking.[17] Empirically, to utilize the moment-fusing flexibility, we further combine the factors obtained from the SR-PCA and ISR-PCA extended moment fusing constructions. The combined factors are either

[17] This adoption helps to reconcile the ISR-PCA reverse eigenvalue ranking (12) with the standard eigenvalue ranking in (16). Specifically, when extending ISR-PCA with pricing restriction, the adjusted transformed covariance matrix is $\frac{1}{T}\widetilde{Y}'\widetilde{Y} + (1+\gamma)\,\overline{Y}'\,\overline{Y}$. Since for ISR-PCA, the transformed volatilities are inverse original SRs, the objective of finding common risk factors that reflect large original SRs is effectuated by changing the sign of the corresponding term $\frac{1}{T}\widetilde{Y}'\widetilde{Y}$ in the above adjusted transformed covariance matrix.

(i) equally weighted, or (ii) optimally weighted

$$\begin{cases} \text{Equally weighted:} & \Pi_{ew} = 0.5\,\Pi_{SR-PCA}(\gamma_e, m_x) + 0.5\,\Pi_{ISR-PCA}(\gamma_e, m_x), \\ \text{Optimally weighted:} & \Pi_{ow} = (1 - w^*)\,\Pi_{SR-PCA}(\gamma_o, m_x) + w^*\,\Pi_{ISR-PCA}(\gamma_o, m_x), \end{cases} \tag{17}$$

wherein the optimal weight w^* is dynamically determined using rolling windows. Sections 4 and 5 below present implementation details and empirical results for FX and equity markets.

4 Empirical Analysis: FX Market

This section presents an empirical analysis of the moment-fusing framework and related pricing models in the FX market. Section 4.1 discusses FX data and sources. Section 4.2 presents basic currency strategies and benchmark FX factors from the perspective of a generic currency denomination (numeraire). Section 4.3 presents in- and out-of-sample empirical results for various moment-fusing constructions and benchmark models in different numeraire currencies. Appendix A.1 details the training procedure of pricing models.

4.1 FX Data

This section describes the exchange rate data and interest differential data required to construct the monthly currency returns. Apart from the U.S. dollar (USD), there are 10 currencies of developed economies in the data: Australian dollar (AUD), Canadian dollar (CAD), Danish krone (DKK), Euro (EUR), Japanese Yen (JPY), New Zealand dollar (NZD), Swedish krona (SEK), Norwegian krone (NOK), Swiss franc (CHF), and British pound sterling (GBP). In the sample, the exchange rates $S_{f/US,t}$ are the price of one unit of currency f in terms of the US dollar at time t (i.e., the exchange rate convention is per unit of the foreign currency).

In the original data set is from the World Market/Refinitiv (WM/R, previous Thompson Reuters), the spot rate (bid and ask) and 1-month forward rate (bid and ask) are all in daily frequencies from December 31, 1984 to October 31, 2025. However, we take the average of the bid and ask rates for both the spot and 1-month forward rates, and take the end-of-month observations throughout the empirical analysis. In addition, before the Euro is introduced in January 1999, we take the German Mark as the proxy for the Euro, and adjust the exchange rate of German Mark by the fixed exchange rate of 1 EUR = 1.95583 German Mark as given by the Deutschbank when Euro is first introduced. Our exchange rate data are therefore monthly from December 1984 to October 2025. The monthly interest rate differential data up to April 2020 is available on Adrien Verdelhan's website. [18] We supplement this data set by employing the forward discount (the one-month forward exchange rate minus the spot exchange rate, in the covered interest parity) to estimate the interest rate differential with respect to the U.S. interest rate for the period from May 2020 to October 2025, relying on the cover interest rate parity (CIP, Footnote 19).

We employ four standard measures to evaluate and quantify a given pricing model, which are (i) the maximum attainable Sharpe ratio, (ii) root-mean-square pricing error, (iii) Gibbons-Ross-Shanken (GRS) test statistic, and (iv) percentage of unexplained (idiosyncratic) return variation. Among these pricing measures, the maximum attainable Sharpe ratio speaks most directly the moment-fusing constructions for two reasons. Namely, SRs are robust to scaling operations of the type (2), and relatedly, moment-fusing factors are constructed to deliver significant factor prices. Reporting and assessing the OOS value of the maximum attainable Sharpe ratio help to mitigate the concern that a model under consideration may over-fitting asset returns by delivering high in-sample SRs. Appendix A.2 presents a description of these pricing measures.

4.2 Currency Returns and FX Strategies

This section discuss currency returns and describes well-known currency strategies and FX pricing factor benchmarks. For concreteness, we first consider USD as the numeraire currency, before adopting a generic numeraire currency H.

USD numeraire: Consider a typical currency strategy borrowing (short) the US dollar at month t to invest (long) in currency i for the period from month t to month $t+1$, at which time the investment is liquidated and proceeds are converted back to the US dollar. Employing CIP, this strategy can be implemented using 1-month forward exchange rate contract. The strategy's realized return, also referred to as currency i's return, in the US dollar numeraire is,[19]

$$CT_{i/US}(t+1) = \ln(S_{t+1}) - \ln(F_t) \equiv s_{t+1} - f_t, \tag{18}$$

[18] http://web.mit.edu/adrienv/www/Data.html, under the heading "Monthly Changes in Exchange Rates and Global Risk Factors".

[19] CIP is a no-arbitrage implication that the currency forward discount $f_t \equiv \frac{S_t}{F_t} - 1$ equals the interest rate differential, $f_t = r_{i,t} - r_{US,t}$, where S_t and F_t are respectively the spot and forward exchange rate between currency i and the US dollar. CIP allows for a substitution of the interest rate differential by the forward discount in the currency return, which results in (18).

where s_{t+1} is the log of the spot rate with respect to the US dollar in month $t+1$. f_t is the log of the 1-month forward exchange rate rate with respect to the US dollar in month t. We compute the currency returns with respect to the US dollar for 10 other currencies in the sample. The currency return series are monthly, spanning January 1985 and October 2025.

We next follow Lustig et al. (2011) to construct two prominent factor in the FX market, namely, the Dollar (denoted as RX) and the High-minus-Low (HML) factors. We sort currency returns (18) based on the contemporaneous forward discount $\ln(S_t/F_t)$ into 5 portfolios, which is equivalent to sorting currencies based on the interest rate differentials of country i with respect to the US, $r_{i,t} - r_{US,t}$, given CIP. The portfolio is rebalanced for each month, so the sorting of currencies into 5 portfolios based on the ranking of the forward discount is for each month. For the month $t+1$, we compute the return of RX and HML as follows,

$$RX(t+1) = \frac{1}{10} \sum_{i=1}^{10} CT_{i/US}(t+1), \tag{19}$$

$$HML(t+1) = \frac{1}{2} \sum_{H_{i,t}=1}^{2} CT_{H_{i,t}/US}(t+1) - \frac{1}{2} \sum_{L_{i,t}=1}^{2} CT_{L_{i,t}/US}(t+1), \tag{20}$$

where $H_{i,t}$ (and $L_{i,t}$) indicates the two currencies with the highest (and lowest) forward discount with respect to the US dollar in month t.

Generic numeraire currency H: We consider the currency return of a generic strategy borrowing currency B and lending other currencies L in the numeraire currency (or, the base currency) H. In terms of the forward and spot rates, this currency return is

$$\begin{aligned} CT^H_{L/B}(t+1) &= \left(s^H_{L,t+1} - s^H_{L,t} + r_{L,t} - r_{H,t}\right) - \left(s^H_{B,t+1} - s^H_{B,t} + r_{B,t} - r_{H,t}\right) \\ &= \left(s^H_{L,t+1} - f^H_{L,t}\right) - \left(s^H_{B,t+1} - f^H_{B,t}\right) = CT^H_{L/H}(t+1) - CT^H_{B/H}(t+1). \end{aligned} \tag{21}$$

where $s^H_{i,t}$ is the log of the spot exchange rate between currency i and the base currency H at time t, and $f^H_{i,t}$ is the log of the forward exchange rate between currency i and the base currency H at time t. Note that the long-short strategy involving currencies L and B in the numeraire currency H can be decomposed into a long-short strategy involving currencies L and H and another one involving currencies H and B. As a result, return $CT^H_{L/B}(t+1)$ (21) is related to a pair of currency i's return $CT_{i/H}(t+1)$ in the numeraire currency H (which are similar to $CT_{i/US}(t+1)$ in (18), with US replaced by H).

To reflect the perspective associated with a new base currency H, we consider the RX strategy as borrowing (short) the home currency and lending (investing in or long) equally other currencies. The return (in the numeraire currency H) of this strategy is

$$RX_H(t+1) = \frac{1}{10}\sum_{i=1}^{10} CT_{i/H}(t+1), \quad \text{where} \quad CT_{i/H}(t+1) = s_{i,t+1}^H - f_{i,t}^H. \tag{22}$$

By construction, RX strategies differ across the respective base currencies as they are home-specific (always borrowing the respective home currencies). As a result, these RX strategy returns do not simply differ from each other by an exchange rate factor between the base currencies. Such a home-specific RX construction allows us to relate their returns to the dominant PC of the corresponding home-specific exchange rates. That is, we aim to examine whether the relationship between RX and the first PC in the US dollar numeraire documented in Lustig et al. (2011) extends to other numeraires.

Similarly, given the perspective associated the base currency H, we sort other currencies based on their forward discount with respect to the base H into 5 portfolios. We consider the HML strategy (from the base H's perspective) as borrowing (short) currencies in the bottom, and lending (long) currencies in the top, portfolios. The return (in the numeraire currency H) of this strategy is

$$HML_{H,t+1} = \frac{1}{2}\sum_{H_{i,t}=1}^{2} CT_{H_{i,t}/H}(t+1) - \frac{1}{2}\sum_{L_{i,t}=1}^{2} CT_{L_{i,t}/H}(t+1), \tag{23}$$

where $H_{i,t}$ (and $L_{i,t}$) indicates the two currencies with the highest (and lowest) forward discount with respect to the base currency H in month t. Since sorting based on the forward discounts is equivalent to sorting based on the interest rate differentials (assuming CIP), the compositions of portfolios underlying the HML strategy in the US dollar base (20) and in the base of currency H (23) are essentially the same.[20] As a result, these HML strategy returns differ only by an exchange rate between the US dollar and currency H.

Moment-fusing constructions in the FX market: The moment-fusing framework is a generic approach to construct pricing factors. For the FX market, we adopt and combine the moment-

[20] In either bases of the US dollar and currency H, the currencies in the top (bottom) portfolio have highest (lowest) interest rates. Therefore, except in periods in which the US dollar or currency H are in these top or bottom portfolios, the long (and short) currencies in the HML strategies are the same from the perspective of the US dollar and H numeraires.

fusion constructions of SR-PCA (7), ISR-PCA (11) with the pricing restrictions from RP-PCA (14) as described in the combined constructions (17), Section 3.3. Our empirical analysis is performed separately for 4 different and representative numeraires of developed currencies in the sample (USD, JPY, AUD, GBP). For each numeraire currency H, we first identify the original asset returns $X_{T\times N}$ in our conceptual analysis (Sections 2 and 3) with the time series of currency returns $\{CT_{i/H}\}$ (22), where i is in the set of $N = 10$ currencies (excluding the current base currency H) in the samples. We then construct the transformed currency returns by incorporating the first moment m_x of the original currency returns $\{CT_{i/H}\}$, $i \in \{1, \ldots, N\}$, using either SR-PCA or ISR-PCA constructions combined with pricing restrictions. As a result, we obtain the associated adjusted covariance matrix $\Sigma_y(m_x)$, whose diagonalization (i.e, the factor analysis, or PCA) leads to the extended moment-fusing (SR-PCA and ISR-PCA) factors (16). To compare with the two FX benchmark factors RX (22) and HML (23), our factor analysis only retains the top two (1st and 2nd) principal factors. We combine the retained principal factors of the same (1st or 2nd) ranking from the extended moment-fusing SR-PCA and ISR-PCA to obtain the corresponding (1st or 2nd) equally and optimally weighted principal factors (17). Finally, to optimize OOS maximum attainable SRs, we incorporate the pricing restrictions into these moment-fusing models, using a training procedure (detailed in Appendix A.1) to determine the pricing restrictions' weight γ dynamically (reported in Figure 1 for the representative USD numeraire).

4.3 Empirical Results in the FX Market

This section reports empirical results and analysis using the FX market data. For references, Table 1 summarizes the nomenclature for various data sets, factor constructions, and pricing measurements employed in our empirical implementation.

[Table 1]

Our empirical results and analysis concern the covariance structure, factor prices (i.e., SRs), and the pricing performance in- and out-of-sample, of the moment-fusing factors and their comparative benchmarks. Among the 4 reported pricing measures, the OOS maximum attainable Sharpe ratio presents a primary assessment for the moment-fusing factors as this measure addresses the factor price, in-sample over-fitting concern, and robustness again scaling operations as discussed earlier.

[Table 2]

Panel A of Table 2 reports, separately for each of the 4 numeraire currencies (USD, JPY, AUD, GBP), the correlations of base-specific RX and HML factors with model-specific top two (1st and 2nd) factors of various models under consideration (standard PCA, RP-PCA, ISR-PCA, and SR-PCA).[21] For all 4 numeraire currencies, the top two factors of RP-PCA practically coincide with those of the standard PCA, as seen in their practically identical correlations with the two FX benchmark factors RX and HML. This indicates a subdued contribution from the pricing restriction term (concerning the first moment) in the RP-PCA objective function (13) for FX data. RX also correlate strongly with the 1st PC for all considered numeraire currencies, extending Lustig et al. (2011)'s finding that RX is basically the level factor to numeraires other than the US dollar. Notably, SR-PCA1 also exhibits similar and strong correlation pattern with the 1st PC, indicating that the top factor of SR-PCA model is also largely the level factor, for all numeraire currencies.

In the numeraire of the US dollar, HML correlates significantly with the 2nd PC, replicating Lustig et al. (2011)'s finding that that HML is basically the slope factor. However, in the numeraire of JPY and AUD, HML correlates moderately with the 2nd PC, indicating that the long-short (currency carry trade) strategy HML is not a level factor universally for all numeraire currencies. Furthermore, the top two factors of ISR-PCA (which are ISR-PCA1 and ISR-PCA2) and the 2nd factor of SR-PCA (which is SR-PCA2) correlate either mildly or little with RX and HML, indicating that these moment-fusing factors differ significantly from the FX benchmark factors RX and HML in all 4 numeraire currencies.

Panel B of Table 2 reports the cross-numeraire correlations of RX (upper half of the panel) and HML (lower half of the panel). Per the discussion below (22), RX strategy (borrowing the home and lending equally all other currencies in the sample) is base-specific (differing principally in their short currencies), and it is also essentially the level factor in that base currency (as seen above, in the Panel A). In the panel, RX cross-numeraire correlations are mild (and mostly negative), indicating that these base-specific level factors (PC1) differ sufficiently. A caveat is that these RX are measured in different base currencies, so their correlations include an inherent exchange rate component between currency bases. The cross-numeraire correlations of HML present a measure for this exchange rate component because HML is largely the same strategy

[21] That is, the top two (1st and 2nd) factors are ranked based on model-specific priority criterion of the model under consideration.

(borrowing low, lending high interest rate currencies) across bases as observed below (23). In the panel, HML's cross-numeraire correlations are significant and positive, affirming significant differences between base-specific level factors.

[Table 3]

Table 3 contains the main results, both in-sample (IS) and out-of-sample (OOS), concerning the pricing performance, quantified by 4 pricing measures (Appendix A.2), of 7 different factor models in the FX market. The models are Lustig et al. (2011)'s RX & HML (22)-(23), PCA, Lettau and Pelger (2020)'s RP-PCA (14), moment-fusing models SR-PCA (7) and ISR-PCA (11), and extended (combined with pricing restrictions) moment-fusing models Π_{ew} (equally weighted) and Π_{ow} (optimally weighted) (17). Among these, RX & HML is a well-known benchmark FX factor pricing model in the related literature. Accordingly, for every model in the table, we retain the top two factors (ranked by the respective priority criterion of that model) for a comparable analysis with the two factors RX and HML. The left panel shows the IS results across the entire sample period, while the right panel shows the OOS results obtained with a dynamic training procedure on the pricing restriction's weight γ (Appendix A.1).

In sample, RX & HML model delivers the highest maximum attainable (annualized) Sharpe ratio of 0.39, 0.43, and 0.34 respectively for 3 numeraire currencies of USD, JPY, and GBP. The extended moment-fusing model Π_{ow} delivers the next highest (and highly comparable) maximum attainable Sharpe ratio of 0.35, 0.45, and 0.34. In the remaining numeraire of AUD, Π_{ow}'s maximum attainable Sharpe ratio of 0.38 outperforms RX & HML's 0.35. The moment-fusing model ISR-PCA delivers significant maximum attainable Sharpe ratios, but also high percentages of unexplained return variation σ_e, reflecting the reverse eigenvalue ranking (i.e., focusing less on the pricing measure of σ_e) in this model. Note that a scaling operation (2) can curb σ_e without changing the identified factors. RP-PCA performs similarly to PCA because as noted in above, pricing restriction has a subdued contribution to the top two principal factors using FX data, producing essentially same top factors in RP-PCA and PCA in sample. Out of sample, moment-fusing factors consistently outperform other models by significant amounts of the highest maximum attainable Sharpe ratio. First, both ISR-PCA and SR-PCA outperforms the benchmark factors RX & HML in all 4 numeraire currencies, and SR-PCA outperforms RP-PCA in 3 numeraire currencies. Second, the extended moment-fusing factors Π_{ow}'s and Π_{ew}'s highest maximum attainable Sharpe ratios further dominate that of moment-fusing factors (ISR-PCA and SR-PCA), PCA and

RP-PCA, and the benchmark RX & HML, in all numeraire currencies. PCA and RP-PCA differ significantly only in the GBP numeraire. These results show the importance and improvement of combining pricing restrictions with the moment-fusing constructions to identify superior pricing factors OOS.

5 Empirical Analysis: Equity Market

This section presents an empirical analysis of the moment-fusing framework and related pricing models in the equity market. Section 5.1 discusses equity data and sources. Section 5.2 presents in- and out-of-sample empirical results for various moment-fusing constructions and benchmark models in various numeraire currencies.

5.1 Equity Data

Our empirical analysis employs 74 and 370 extreme-decile anomaly portfolios (denoted respectively as LP74 and LP370) from Lettau and Pelger (2020) and sampled at monthly frequency. They select 37 out of 50 characteristics that are available as of November 1963, and sort returns into 10 portfolios based on the characteristics deciles.[22] While the sourced data (Footnote 22) is available up to December 2019, we follow the source's definitions of anomalies to replicate and extend the anomaly portfolios up to December 2024. We also obtain and employ the monthly returns of the Fama-French 25 size-B/M sorted portfolios (denoted as FF25).[23] As a result, in our empirical analysis, the sample period for these three data sets is from November 1963 to December 2024 with no missing values of the these portfolio returns. The standard deviations and kurtosis of portfolio returns are within the normal range. For LP74, the average standard deviations across 74 portfolio monthly returns is 5%, and the average kurtosis is 3. For LP370, the average standard deviations across 370 portfolio monthly returns is 5%, and the average kurtosis is 3. Therefore, we do not further winsorize or truncate the portfolio returns.

Our main equity test assets are LP74 and LP370 due to the sufficiently large number of equity portfolios and diverse characteristics underlying those portfolios. For robustness, we further include FF25 as test assets. Among the 37 characteristics underlying LP74 and LP370, 14

[22] The characteristics are available at https://www.serhiykozak.com/data, under the heading "Portfolio Sorts", and discussed in Kozak et al. (2018).

[23] Fama-French 25 size-B/M sorted portfolios are from Ken French's website http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data-library.html

are constructed from price measures. Other 21 characteristics are constructed from accounting measures, of which 9 concern valuation and profitability measures. The list of these characteristics is presented in Appendix A.3. The moment-fusing constructions are similar to those for the FX market. In particular, to optimize OOS maximum attainable SRs, we incorporate the pricing restrictions into these moment-fusing constructions, using a training procedure (Appendix A.1) to determine the pricing restrictions' weight γ dynamically (reported in Figure 2 for the representative LP74 portfolios).

5.2 Empirical Results in the Equity Market

This section reports empirical results and analysis using the equity market data. Our empirical results and analysis concern the pricing performance of the moment-fusing factors and their comparative benchmark models separately for three data sets (LP74, LP370, FF25) in-sample and out-of-sample. Similar to the empirical analysis using FX data above, due to the nature of the moment-fusing constructions, the OOS maximum attainable Sharpe ratio remains a primary pricing measure for the moment-fusing factors using equity data. For robustness, we also report 3 other standard pricing measures for all pricing models (Appendix A.2). The the nomenclature is in Table 1.

[Table 4]

Table 4 reports the pricing performance, quantified by 4 pricing measures (Appendix A.2), of 6 different factor models using 74 extreme-decile anomaly portfolios. The models are PCA, Lettau and Pelger (2020)'s RP-PCA (14), moment-fusing models SR-PCA (7) and ISR-PCA (11), and extended (combined with pricing restrictions) moment-fusing models Π_{ew} (equally weighted) and Π_{ow} (optimally weighted) (17). Among these, RP-PCA is a prominent factor pricing benchmark model that incorporates both first and second moments of returns. Accordingly, for every model in the table, we retain $K = 3$ and $K = 5$ factors (ranked by the respective priority criterion of that model) for a comparable analysis. The left panel shows the IS results across the entire sample period, while the right panel shows the OOS results obtained with a dynamic training procedure on the pricing restriction's weight γ (Appendix A.1).

In sample, for $K = 3$ factors, the moment-fusing models SR-PCA, Π_{ow}, Π_{ew} and RP-PCA in order deliver 4 highest maximum attainable (monthly) Sharpe ratio of 0.45, 0.45, 0.44 and 0.37

respectively. For $K = 5$ factors, the same ranking holds for these 4 models but with slightly improved Sharpe ratio of 0.49, 0.49, 0.47 and 0.45. Note that these Sharpe ratios (after being converted to annualized values) are significantly higher than those obtained using FX data (Section 4.3) in part because a larger number of returns in the equity data sample offers a higher degree of diversification. Out of sample, for $K = 3$ factors, SR-PCA, Π_{ow}, and RP-PCA deliver similar maximum attainable (monthly) Sharpe ratio of 0.09, 0.09, and 0.10 respectively. For $K = 5$, these models deliver 0.55, 0.55, and 0.45 respectively. The OOS under-performance of Π_{ew} compared to Π_{ow} (for both $K = 3$ and $K = 5$) shows the robustness of optimally choosing the weight between components in the extended moment-fusing model (17). The moment-fusing model ISR-PCA delivers high maximum attainable Sharpe ratio (0.41 for $K = 3$ and 0.48 for $K = 5$) but also the highest percentage of unexplained return variation σ_e, reflecting the reverse eigenvalue ranking in this model.

[Table 5]

Table 5 reports the pricing performance of the same models using 370 extreme-decile anomaly portfolios. In sample, for both $K = 3$ and $K = 5$ factors, the moment-fusing models Π_{ow}, Π_{ew}, SR-PCA, and ISR-PCA deliver 4 highest maximum attainable (monthly) Sharpe ratio. Out of sample, for $K = 3$ factors, Π_{ow} and PCA deliver highest maximum attainable (monthly) Sharpe ratio (0.21, and 0.22 respectively), next performing models are RP-PCA and SR-PCA (delivering 0.19, and 0.16 respectively). For $K = 5$, Π_{ow}, SR-PCA and RP-PCA are three top models, all delivering the highest maximum attainable (monthly) Sharpe ratio of 0.37, while PCA delivers 0.26.

[Table 6]

Table 6 reports the pricing performance of the same models using the Fama-French 25 size-B/M sorted portfolios, which has a significantly smaller sample size compared to LP74 and LP370. In sample, the moment-fusing models Π_{ow}, SR-PCA, and RP-PCA are top performing models in terms of delivering the highest maximum attainable (monthly) Sharpe ratio of respectively 0.25, 0.25, and 0.26 (for $K = 3$), and 0.37, 0.37, and 0.36 (for $K = 5$). Out of sample, these same models deliver the maximum attainable Sharpe ratio of 0.24, 0.24, and 0.25 (for $K = 3$), and 0.28, 0.28, and 0.34 (for $K = 5$), indicating the out-performance of RP-PCA when we retain 5 factors.

In summary, the moment-fusing factor construction extended with pricing restrictions and optimal weight Π_{ow} is the outperforming pricing model OOS for consistently for all equity data sets, different numbers of retained factors, in terms of delivering maximum attainable Sharpe ratio and other pricing measures (with the exception of a 5-factor setting using FF25 data set, in which RP-PCA is the outperforming model). Next, the moment-fusing factor model $SR-PCA$ also outperforms OOS for larger equity data sets of LP74 and LP370, specially when we retain $K = 5$ factors. These findings show important improvements in pricing factors whose constructions are fused with original return moments combined with the imposition of pricing restrictions.

6 Conclusions

The current paper introduces the moment-fusing framework to identify important common risk factors that price asset returns in the cross section and out of sample. The framework consists of two stages. In the first stage, we construct an equivalent set of transformed returns whose covariance structure is fused with information of the first (and possibly higher) moment of original returns. In the second stage, we implement the covariance-based factor analysis on the transformed returns. This approach obtains the principal pricing factors are informed by the risk premia and Sharpe ratios of original returns while also utilizing the elegance and power of the PCA methodology.

Using FX and equity data, our empirical analysis provides evidence for the out-performance of the moment-fusing models as quantified by the maximum attainable Sharpe ratios and other pricing measures. The moment-fusing constructions are flexible and can also be combined with pricing restrictions to enhance their pricing performance. While the current paper focuses on optimizing factor prices, the moment-fusing framework is general and adaptable to optimizing other moments of pricing factors.

References

Barillas, F., and J. Shanken, 2017, Which Alpha?, *Review of Financial Studies* 30, 1316–1338.

Chamberlain, G., and M Roth, 1983, Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets., *Econometrica* 51, 1281–1304.

Cochrane, J., 2011, Presidential Address: Discount Rates., *Journal of Finance* 66, 1047–1108.

Connor, G., and R Korajczyk, 1986, Performance Measurement with the Arbitrage Pricing Theory: A New Framework for Analysis., *Journal of Financial Economics* 15, 373–394.

Connor, G., and R Korajczyk, 1988, Risk and Return in an Equilibrium APT: Application of a New Test Methodology., *Journal of Financial Economics* 21, 255–289.

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

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

Giglio, S., D. Xiu, and D. Zhang, 2025, Test Assets and Weak Factors, *Journal of Finance* 80, 259–319.

Harvey, C., Y. Liu, and H. Zhu, 2016, ... and the Cross-Section of Expected Returns, *Review of Financial Studies* 29, 5–68.

Huang, D., F. Jiang, K. Li, G. Tong, and G. Zhou, 2022, Scaled PCA: : A New Approach to Dimension Reduction, *Management Science* 68, 1678–1695.

Kan, R., X. Wang, and X Zheng, 2024, In-sample and Out-of-sample Sharpe Ratios of Multi-factor Asset Pricing Models., *Journal of Financial Economics* 155.

Kelly, B., S. Pruitt, and Y. Su, 2019, Characteristics Are Covariances: A Unified Model of Risk and Return, *Journal of Financial Economics* 134, 501–524.

Kozak, S., S. Nagel, and S. Santosh, 2018, Interpreting Factor Models, *Journal of Finance* 73, 1183–1223.

Lettau, M., and M. Pelger, 2020, Factors That Fit the Time Series and Cross-Section of Stock Returns, *Review of Financial Studies* 33, 2274–2325.

Litterman, R., and J Scheinkman, 1991, Common Factors Affecting Bond Returns., *Journal of Fixed Income* 1, 54–61.

Lustig, H., N. Roussanov, and A. Verdelhan, 2011, Common Risk Factors in Currency Returns, *Review of Financial Studies* 24, 3731–3777.

Markowitz, H., 1952, Portfolio Selection, *Journal of Finance* 7, 77–91.

Maurer, T., T. Tô, and N-K. Tran, 2019, Pricing Risks Across Currency Denominations, *Management Science* 65, 5308–5336.

Ross, S., 1976, The Arbitrage Theory of Capital Asset Pricing., *Journal of Economic Theory* 13, 341–360.

Tables and Figures

Table 1: **Nomenclature**

This table outlines the nomenclature of each method in the paper.

Panel A	**Factor Models**
PCA	Principal component analysis without adding pricing restrictions.
RX& HML	RX is the dollar factor and HML is the carry trade factor, see Lustig et al. (2011).
RP-PCA	Principal component analysis using the pricing restriction (γ). that leads to the highest OOS Sharpe ratio
ISR-PCA	Equalize mean asset returns and run PCA, sort factors by eigenvalues from smallest to largest, then apply pricing restrictions.
SR-PCA	Run PCA in the scaled asset returns, then apply pricing restrictions.
Π_{ew}	Equal-weighted portfolio constructed by ISR-PCA and SR-PCA.
Π_{ow}	The weight assigned to the portfolio by ISR-PCA and SR-PCA with the highest in-sample SR.
Panel B	**Estimates**
No Adjust	The covariance matrix is estimated from the historical mean return
Panel C	**Data Sets**
LP74	The 74 extreme decile anomaly portfolios analyzed by Lettau and Pelger (2020)
LP370	370 anomaly portfolios analyzed by Lettau and Pelger (2020)
FF25	Fama-French 25 size-B/M sorted portfolios
Panel D	**Pricing Measures**
SR	The maximum Sharpe ratio attained with the factors
$RMSE_{\alpha}$	The square root of the average pricing error
$\overline{\sigma}_e$	The percentage of unexplained return variation
GRS	GRS test statistic

Table 2: **Correlation of Foreign Exchange Factors**

This table reports the correlation coefficients of the top two factors in different FX pricing models. The factors are constructed at monthly frequency in four numeraire currencies (USD, JPY, AUD, and GBP) from the data set of 11 developed currencies spanning the period of 1985:1 – 2025:10. RX and HML are base-specific factors from Lustig et al. (2011). Panel A reports the correlations, separately for each numeraire currency, for the top two factors of 5 models with base-specific factors RX and HML. Panels B reports the cross-numeraire correlations respectively for base-specific factors RX and HML after converting their numeraire currencies into the US dollar, computed by dividing the returns of local currency by the exchange rate with respect to the US dollar from that month.

Panel A					
	RX-USD	HML-USD		RX-JPY	HML-JPY
PC1-USD	0.99	0.19	PC1-JPY	0.99	0.40
PC2-USD	0.02	0.69	PC2-JPY	−0.02	0.31
RP-PCA1-USD	0.99	0.19	RP-PCA1-JPY	0.99	0.40
RP-PCA2-USD	−0.02	0.69	RP-PCA2-JPY	−0.03	0.31
ISR-PCA1-USD	0.85	−0.02	ISR-PCA1-JPY	0.83	0.18
ISR-PCA2-USD	0.61	0.41	ISR-PCA2-JPY	0.59	0.50
SR-PCA1-USD	0.99	0.24	SR-PCA1-JPY	0.99	0.40
SR-PCA2-USD	−0.30	0.53	SR-PCA2-JPY	0.16	0.47
	RX-AUD	HML-AUD		RX-GBP	HML-GBP
PC1-AUD	0.99	−0.39	PC1-GBP	0.99	0.09
PC2-AUD	0.01	−0.23	PC2-GBP	−0.06	0.70
RP-PCA1-AUD	0.99	−0.39	RP-PCA1-GBP	0.99	0.08
RP-PCA2-AUD	−0.01	−0.24	RP-PCA2-GBP	−0.05	0.73
ISR-PCA1-AUD	0.92	−0.39	ISR-PCA1-GBP	0.25	−0.35
ISR-PCA2-AUD	−0.26	0.26	ISR-PCA2-GBP	−0.40	0.05
SR-PCA1-AUD	0.99	−0.37	SR-PCA1-GBP	0.98	−0.03
SR-PCA2-AUD	−0.42	0.46	SR-PCA2-GBP	0.39	0.70

Panel B	RX-USD	RX-JPY	RX-AUD	RX-GBP
RX-USD	1			
RX-JPY	0.17	1		
RX-AUD	−0.05	−0.31	1	
RX-GBP	−0.02	−0.14	−0.24	1
	HML-USD	HML-JPY	HML-AUD	HML-GBP
HML-USD	1			
HML-JPY	0.83	1		
HML-AUD	0.86	0.61	1	
HML-GBP	0.80	0.54	0.80	1

Table 3: **Currency Portfolios Results**

This table reports the point estimates of four different pricing measures for various two-factor FX pricing models. The individual factors are constructed at monthly frequency in four numeraire currencies (USD, JPY, AUD, GBP) from the data set of 11 developed currencies spanning the period of 1985:1 - 2025:10. All point estimates are annualized, p is the p-value of the GRS test statistic. The description of the nomenclature refers to Table 1. Red font indicates the literature's main benchmark model. Blue font indicates the best model in the Table based on the measure of max Sharpe ratio.

	In-Sample						**Out-of-Sample**				
	SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p		SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p
Panel A: USD											
RX & HML	0.39	0.01	30.68	0.25	0.99	RX & HML	0.09	0.01	23.69	0.40	0.91
PCA	0.24	0.01	23.88	0.23	0.99	PCA	0.08	0.01	24.42	0.42	0.94
RP-PCA	0.24	0.01	23.88	0.22	0.99	RP-PCA	0.09	0.01	24.41	0.42	0.94
ISR-PCA	0.35	0.01	46.20	0.50	0.89	ISR-PCA	0.28	0.01	34.81	0.83	0.60
SR-PCA	0.27	0.01	24.50	0.22	0.99	SR-PCA	0.09	0.01	24.96	0.45	0.92
Π_{ew}	0.32	0.01	26.56	0.21	1.00	Π_{ew}	0.18	0.01	24.95	0.46	0.92
Π_{ow}	0.35	0.01	46.20	0.50	0.89	Π_{ow}	0.28	0.01	34.81	0.83	0.60
Panel B: JPY											
RX & HML	0.43	0.01	24.40	0.26	0.99	RX & HML	0.18	0.01	17.14	0.57	0.84
PCA	0.21	0.01	16.85	0.33	0.97	PCA	0.31	0.01	17.92	0.34	0.97
RP-PCA	0.21	0.01	16.86	0.33	0.97	RP-PCA	0.32	0.01	17.92	0.34	0.97
ISR-PCA	0.40	0.01	47.84	0.68	0.74	ISR-PCA	0.43	0.01	20.39	0.48	0.90
SR-PCA	0.25	0.01	17.39	0.31	0.98	SR-PCA	0.36	0.01	17.57	0.32	0.98
Π_{ew}	0.29	0.01	19.01	0.32	0.98	Π_{ew}	0.38	0.01	17.49	0.33	0.97
Π_{ow}	0.42	0.01	70.09	0.32	0.98	Π_{ow}	0.61	0.02	30.75	1.22	0.27
Panel C: AUD											
RX & HML	0.35	0.01	27.41	0.43	0.92	RX & HML	0.01	0.01	31.68	0.38	0.95
PCA	0.21	0.01	20.72	0.30	0.98	PCA	0.15	0.01	27.88	0.40	0.95
RP-PCA	0.23	0.01	20.74	0.29	0.98	RP-PCA	0.18	0.01	27.92	0.40	0.94
ISR-PCA	0.12	0.01	42.44	0.59	0.83	ISR-PCA	0.08	0.01	30.69	0.38	0.95
SR-PCA	0.37	0.00	25.83	0.08	1.00	SR-PCA	0.25	0.01	32.68	0.45	0.92
Π_{ew}	0.36	0.00	25.93	0.10	1.00	Π_{ew}	0.60	0.01	34.50	0.62	0.79
Π_{ow}	0.38	0.00	25.47	0.07	1.00	Π_{ow}	0.49	0.01	33.83	0.56	0.85
Panel D: GBP											
RX & HML	0.34	0.01	39.04	0.31	0.98	RX & HML	0.11	0.01	33.82	0.37	0.96
PCA	0.17	0.01	32.27	0.46	0.91	PCA	0.13	0.01	38.73	0.40	0.94
RP-PCA	0.20	0.01	32.32	0.42	0.94	RP-PCA	0.27	0.01	38.79	0.42	0.94
ISR-PCA	0.16	0.01	75.18	0.62	0.79	ISR-PCA	0.19	0.01	43.11	0.82	0.60
SR-PCA	0.33	0.01	35.68	0.16	1.00	SR-PCA	0.23	0.01	38.45	0.41	0.94
Π_{ew}	0.33	0.01	38.70	0.23	0.99	Π_{ew}	0.86	0.01	38.23	0.37	0.96
Π_{ow}	0.34	0.01	36.05	0.20	1.00	Π_{ow}	0.61	0.01	38.22	0.37	0.96

Table 4: **Lettau-Pelger 74 Portfolio Results**

This table reports the point estimates of four different pricing measures for factor models using the 74 extreme decile anomaly portfolios (Lettau and Pelger (2020)). The sample period is 1963:11- 2024:12 at monthly frequency. All point estimates are at monthly level, p is the p-value of the GRS test statistic. The pricing restrictions term γ applied in this table is the γ that generates the best SR. The description of the nomenclature refers to Table 1. Red font indicates the literature's main benchmark model. Blue font indicates the best model in the Table based on the measure of max Sharpe ratio.

	In-Sample					**Out-of-Sample**					
	SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p		SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p
Panel A: 3-Factor											
PCA	0.18	0.30	16.37	6.30	0	PCA	0.15	0.31	18.34	5.63	0
RP-PCA	0.37	0.27	16.70	4.95	0	RP-PCA	0.10	0.32	18.34	5.79	0
ISR-PCA	0.32	0.57	93.89	5.35	0	ISR-PCA	0.41	0.27	20.59	4.41	0
SR-PCA	0.45	0.27	17.14	4.23	0	SR-PCA	0.09	0.34	18.40	5.49	0
Π_{ew}	0.44	0.26	17.29	4.23	0	Π_{ew}	0.01	0.36	18.29	5.64	0
Π_{ow}	0.45	0.27	17.14	4.23	0	Π_{ow}	0.09	0.34	18.40	5.49	0
Panel B: 5-Factor											
PCA	0.28	0.24	12.39	5.65	0	PCA	0.30	0.24	13.55	5.14	0
RP-PCA	0.45	0.23	12.52	4.18	0	RP-PCA	0.45	0.24	13.57	5.05	0
ISR-PCA	0.35	0.62	76.55	5.13	0	ISR-PCA	0.48	0.23	16.31	4.49	0
SR-PCA	0.49	0.24	12.86	3.80	0	SR-PCA	0.55	0.23	13.71	5.10	0
Π_{ew}	0.47	0.23	12.99	3.95	0	Π_{ew}	0.29	0.23	13.75	4.90	0
Π_{ow}	0.49	0.24	12.86	3.80	0	Π_{ow}	0.55	0.23	13.72	5.11	0

Table 5: **Lettau-Pelger 370 Portfolio Results**

This table reports the point estimates of four different pricing measures for factor models using the 370 anomaly portfolios (Lettau and Pelger (2020)). The sample period is 1963:11- 2024:12 at monthly frequency. All point estimates are at monthly level, p is the p-value of the GRS test statistic. The pricing restrictions term γ applied in this table is the γ that generates the best SR. The description of the nomenclature refers to Table 1. Red font indicates the literature's main benchmark model. Blue font indicates the best model in the Table based on the measure of max Sharpe ratio.

	In-Sample						**Out-of-Sample**				
	SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p		SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p
Panel A: 3-Factor											
PCA	0.18	0.17	14.45	2.12	0	PCA	0.22	0.17	16.14	3.13	0
RP-PCA	0.24	0.17	14.50	2.05	0	RP-PCA	0.19	0.17	16.16	3.13	0
ISR-PCA	0.36	0.53	96.91	1.87	0	ISR-PCA	0.13	0.21	18.56	2.82	0
SR-PCA	0.40	0.18	15.19	1.81	0	SR-PCA	0.16	0.18	16.28	3.04	0
Π_{ew}	0.40	0.19	15.20	1.79	0	Π_{ew}	0.15	0.17	16.28	3.02	0
Π_{ow}	0.40	0.18	15.19	1.81	0	Π_{ow}	0.21	0.17	16.29	3.04	0
Panel B: 5-Factor											
PCA	0.23	0.15	12.08	2.05	0	PCA	0.26	0.15	13.02	3.07	0
RP-PCA	0.34	0.15	12.12	1.89	0	RP-PCA	0.37	0.15	13.03	3.06	0
ISR-PCA	0.37	0.47	92.40	1.86	0	ISR-PCA	0.16	0.18	15.10	2.84	0
SR-PCA	0.44	0.17	12.28	1.74	0	SR-PCA	0.37	0.15	12.95	3.03	0
Π_{ew}	0.44	0.17	12.31	1.72	0	Π_{ew}	0.20	0.15	12.95	3.05	0
Π_{ow}	0.44	0.17	12.28	1.74	0	Π_{ow}	0.37	0.15	12.95	3.01	0

Table 6: **Fama-French 25 Portfolio Results**

This table reports the point estimates of four different pricing measures for factor models using the Fama-French 25 size-B/M sorted portfolios. The sample period is 1963:11- 2024:12 at monthly frequency. All point estimates are at monthly level, p is the p-value of the GRS test statistic. The pricing restrictions term γ applied in this table is the γ that generates the best SR. The description of the nomenclature refers to Table 1. Red font indicates the literature's main benchmark model. Blue font indicates the best model in the Table based on the measure of max Sharpe ratio.

	In-Sample						**Out-of-Sample**				
	SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p		SR	$RMSE_\alpha$	$\overline{\sigma}_e$	GRS	p
Panel A: 3-Factor											
PCA	0.25	0.14	6.62	4.05	0	PCA	0.25	0.17	7.20	6.02	0
RP-PCA	0.26	0.14	6.63	3.93	0	RP-PCA	0.25	0.17	7.21	6.04	0
ISR-PCA	0.10	1.05	96.69	5.85	0	ISR-PCA	0.23	0.29	19.58	6.03	0
SR-PCA	0.25	0.15	6.71	4.06	0	SR-PCA	0.24	0.17	7.24	6.35	0
Π_{ew}	0.25	0.15	7.06	4.08	0	Π_{ew}	0.24	0.17	7.25	6.36	0
Π_{ow}	0.25	0.15	6.71	4.06	0	Π_{ow}	0.24	0.17	7.24	6.35	0
Panel B: 5-Factor											
PCA	0.28	0.12	4.66	3.57	0	PCA	0.28	0.15	4.92	5.75	0
RP-PCA	0.36	0.10	4.72	2.16	0	RP-PCA	0.34	0.14	4.92	5.22	0
ISR-PCA	0.18	0.79	86.62	5.06	0	ISR-PCA	0.27	0.26	15.76	6.86	0
SR-PCA	0.37	0.10	4.81	2.09	0	SR-PCA	0.28	0.16	4.98	5.97	0
Π_{ew}	0.33	0.10	5.36	2.79	0	Π_{ew}	0.28	0.16	5.04	6.14	0
Π_{ow}	0.37	0.10	4.81	2.09	0	Π_{ow}	0.28	0.16	4.98	5.97	0

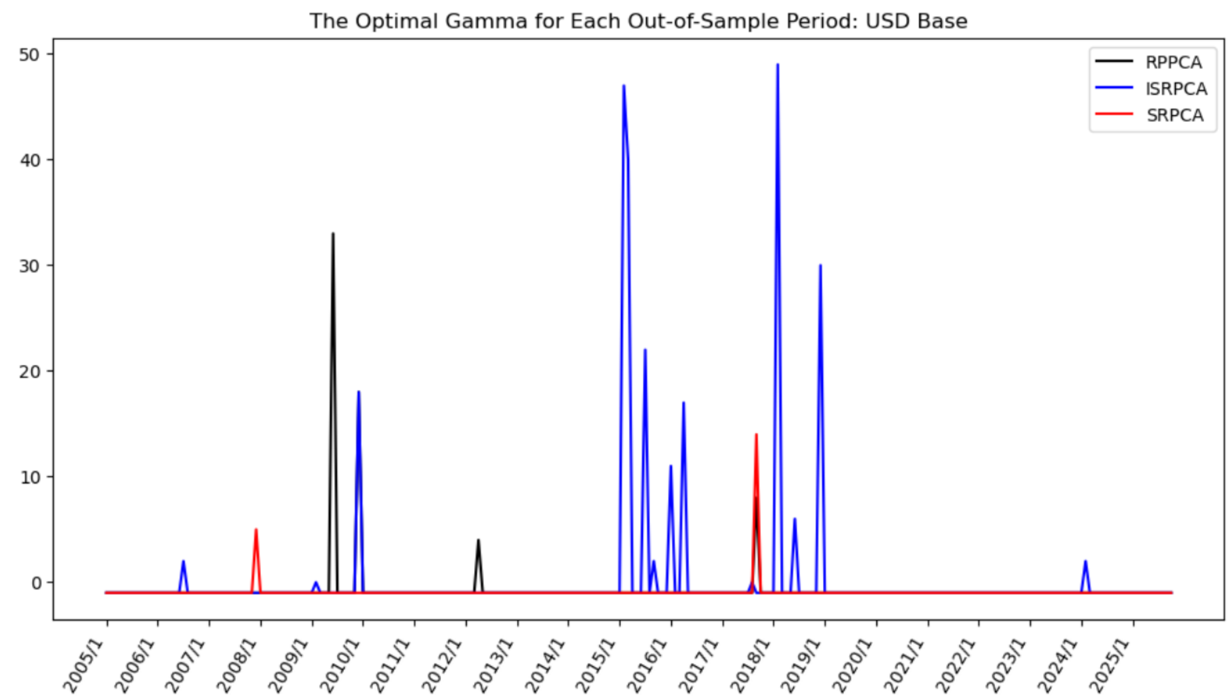

Figure 1: **The Time Series of the Optimal γ of the Currency Carry Trade Return Data: USD Base**. This plot shows the time series of the γ in each out-of-sample period that leads to the optimal Sharpe ratios for the USD base. The γ is based on the covariance matrix adjusted for their eigenvalues.

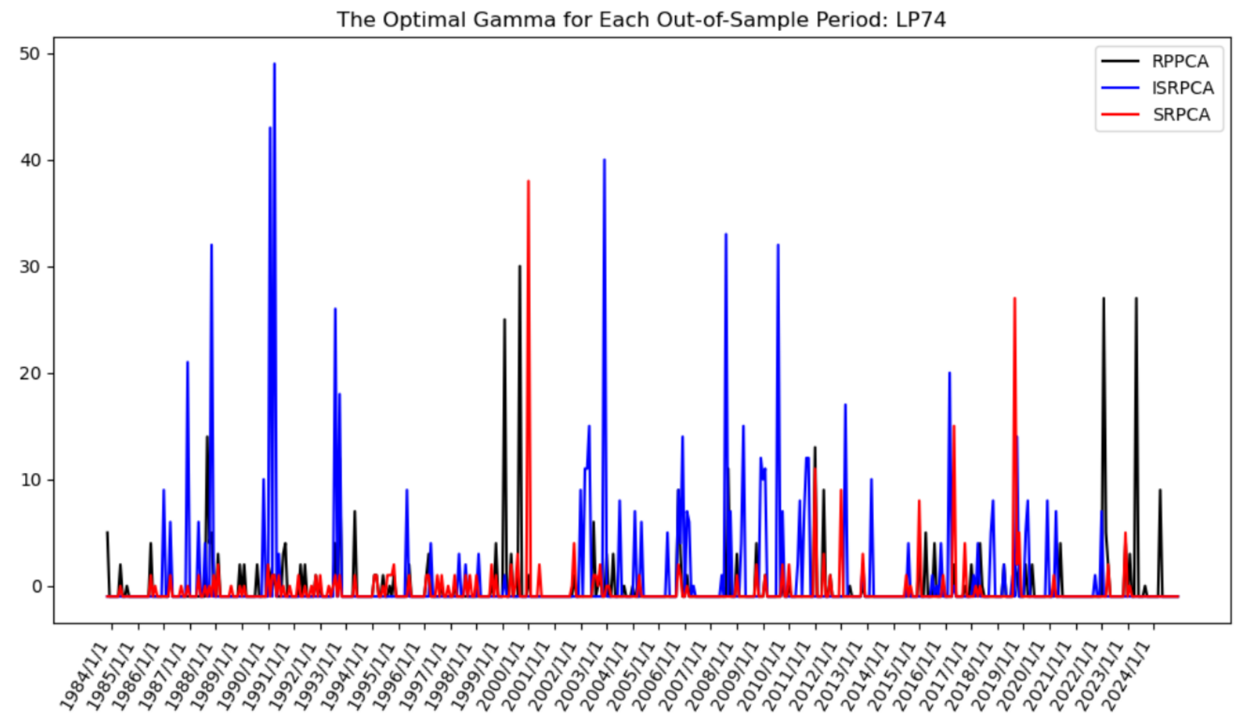

Figure 2: **The Time Series of the Optimal γ of the Stock Portfolio Return Data: LP74**. This plot shows the time series of the γ in each out-of-sample period that leads to the optimal Sharpe ratios for the LP74 test asset. The γ is based on the covariance matrix adjusted for their eigenvalues.

Online Appendices

Appendix A contains further data and empirical details, Appendix A.1 summarizes the training steps for pricing models, Appendix A.2 discusses four pricing measures employed in the paper, Appendix A.3 lists characteristics concerning equity portfolios.

A Further Data and Empirical Details

A.1 The Training Procedure for the Pricing Restriction Parameter

We follow the approach of Lettau and Pelger (2020) to impose the pricing restriction parameter γ to regularize the cross-sectional pricing errors and integrate it with the standard principal component analysis. The objective function of this optimization is specified in (13). We further employs the following training procedure to determine γ dynamically. Given a data set of generic traded asset returns, the procedure trains the optimal γ to obtain the optimal out-of-sample Sharpe ratios based on K factors. We employ $K = 5$ in the equity return data and $K = 2$ in the FX data are as follows.

1. We begin with an initial 240-month training window ($t = 0$). In the initial We use the information from the initial 240-month training window to forecast returns for month $t = 1$ across different numerical values inserted into the pricing restriction parameter γ.[24] That is, we form the out-of-sample factor in month 1 by

$$F_{OOS}(t=1) = X(t=1)\Lambda(t \rightarrow 0, \gamma_{t=1}) \tag{24}$$

   We then record the out-of-sample Sharpe ratio of the K-factor model, $SR_{OOS,t=1}$, from the out-of-sample factor $F_{OOS}(t = 1)$ and identify the optimal $\gamma_{t=1}$ that maximizes $SR_{OOS,t=1}$.

2. We expand the training window by one month (now 241 months) and repeat the process to forecast returns for $t = 2$. That is, we form the out-of-sample factor in month 2 by

$$F_{OOS}(t=2) = X(t=2)\Lambda(t \rightarrow 1, \gamma_{t=2}) \tag{25}$$

[24] The pricing-restriction parameter is searched over $[-1, 15]$ for LP74, $[-1, 10]$ for LP370, and $[-1, 50]$ for FF25 and all four FX data sets. The difference in search ranges is due to different computational constraints in different data sets. Moreover, for all data sets and all methods, γ at each time point rarely exceeds 20.

We compute $SR_{OOS,t=2}$ from the out-of-sample factor $F_{OOS}(t = 2)$ and select the corresponding optimal $\gamma_{t=2}$.

3. We repeat steps 1 and 2 as above until the end of the observations in the sample. In this way, we obtain a time series of γ with each observation the optimal γ that generates the highest out-of-sample Sharpe ratios for the K-factor model at each time point and a time series of out-of-sample Sharpe ratios SR_{OOS}.

To illustrate how the pricing restriction evolves across the time, we plot the time series of the optimal γ for the 2-factor models of RP-PCA, ISR-PCA, and SR-PCA in the carry trade return data without adjusting the eigenvalues are in Figure 1. We plot of the optimal γ for the 5-factor models of RP-PCA, ISR-PCA, and SR-PCA in the equity data without adjusting the eigenvalues in Figure 2.

From the time series plot of the pricing restriction parameter γ in the carry trade portfolio data, we observe that for most of the time, the optimal γ is chosen to be -1, meaning that there is not any pricing restriction incurred in the covariance matrix when estimating the out-of-sample Sharpe ratio during the training process. However, for ISR-PCA, we observe that the optimal γ spikes around the middle of the sample period, which indicates that adding the pricing restriction helps to partially stabilize the estimation of out-of-sample Sharpe ratios.

From the time series plot of the pricing restriction parameter γ in the carry trade portfolio data, we observe that for most of the time, the optimal γ is also chosen to be -1. However, compared to the case in the carry trade return data, there are much more number of months in which the optimal γ spikes up than the case in the carry trade return data. One reason is that there might be much more noise in the stock factor portfolio return data, so adding the pricing restriction helps to stabilize the estimation of the out-of-sample Sharpe ratio more than in the carry trade return data.

We also describe the technique to construct the portfolio that combines the ISR-PCA and SR-PCA methods, such that the combined portfolio, if finding the optimal weight, should deliver a higher Sharpe ratio than the individual portfolio. We apply the weight of w_1 to ISR-PCA factor and $1-w_1$ to SR-PCA factor to combine them into a new factor, where we construct ISR-PCA and SR-PCA factors with the γ in the procedures above. In particular,

- Π_{ew}: Apply $w_1 = 0.5$ to both the ISR-PCA and SR-PCA factors.

- Π_{ow}: Apply the weight w_1 to the combined factor that earns the maximum in-sample Sharpe ratios of $K = 5$ in equity return data and $K = 2$ in the FX data.

The optimal weight applied to ISR-PCA in each data set is as follows: USD Base is 1. JPY Base is 1.05. AUD Base is 0.4. GBP Base is 0.35. LP74 is 0. LP370 is 0.02. FF25 is 0. These optimal weights indicate that in most data sets, the training algorithm determines that the best combination of model that generates the highest in-sample Sharpe ratio is either pure ISR-PCA (USD, JPY) or pure SR-PCA (LP74, LP370, FF25). Only in AUD or GBP does the algorithm determine the optimal combination of the two models is different from the pure ISR-PCA and SR-PCA. Finally, this optimal weight is determined using the covariance matrix without further adjusting for their eigenvalues.

The **in-sample Sharpe ratios** of the combined portfolio from the ISR-PCA and SR-PCA factor are computed as follows,

1. Denote $\gamma_{ISR-PCA}$ as the γ applied to ISR-PCA and γ_{SR-PCA} as the γ applied to SR-PCA. The in-sample results are computed using the full sample. A single value of γ searched over the interval $[-1, 50]$ is selected and then used to estimate the factors that maximize the in-sample Sharpe ratio across the entire sample period.

2. The equally-weighted combined factor is by applying $w_1 = 0.5$ to ISR-PCA and $w_2 = 0.5$ to SR-PCA. The equal-weight factor is

$$\Pi_{ew} = 0.5F_{ISR-PCA} + 0.5F_{SR-PCA} \tag{26}$$

To search for the weight w_1^* that leads to the optimal in-sample Sharpe ratio of the combined factor, we denote $w_1 \equiv \Sigma_{ISR}^{-1}\boldsymbol{\mu}$ as the optimal weight applied to ISR-PCA, and insert the numerical values over the range of -2 and 2 and determine which numerical value inserted into w_1 generates the best in-sample Sharpe ratio. That optimal-weight factor is

$$\Pi_{ow} = w_1^*F_{ISR-PCA} + (1 - w_1^*)F_{SR-PCA} \tag{27}$$

To compute the **out-of-sample Sharpe ratios** from Π_{ew} and Π_{ow} generated above, we implement the steps that produce the results as follows,

1. We start with the initial window of 240 months, denote month 240 as time 0 have a loading

matrix $\Lambda_{OW,0}$ (in the optimal weight) and or $\Lambda_{EW,0}$ (in the equal weight), predict the return in the month 1 as the first out-of-sample period. We have the first observation of the out-of-sample Sharpe ratio.

2. The length of the training window becomes 241 months, and we now predict the return in month 2 with $\Lambda_{OW,1}$ (in the optimal weight) and or $\Lambda_{EW,1}$ (in the equal weight). We have the second observation of the out-of-sample Sharpe ratio. In these steps, we form the equal-weighted and optimal-weighted combined factors as

$$\begin{aligned}\Pi_{ew}(w_1,\gamma) =& 0.5X\Lambda_{ISR}(\gamma_{EW}) + 0.5X\Lambda_{SR}(\gamma_{EW})\\ \Pi_{ow}(w_1,\gamma) =& w_1X\Lambda_{ISR}(\gamma_{OW}) + (1-w_1)X\Lambda_{SR}(\gamma_{OW})\end{aligned} \tag{28}$$

   Where γ_{EW} and γ_{OW} are the optimal γ that should apply to this equal-weighted or optimal-weighted combined factors, which are different from the γ applied to the individual factors of ISR-PCA or SR-PCA.

3. We repeat this process until the end of the observation. We compute the time-series average of the out-of-sample Sharpe ratio as our out-of-sample results.

A.2 Pricing Measures

For a self-contained connection between actual test statistics and the conceptual description of the moment-fusing framework, we briefly describe the SR estimate and basic features of pricing measures employed in our empirical analysis. We consider a given pricing model G_K characterized by K traded pricing factors $\{\Pi_i\}_{i=1}^K$ and a set of N test assets $\{X_n\}_{n=1}^N$. Factors are linear in asset returns, $\Pi = XW$ (with the demeaned version, $\widetilde{\Pi} = \widetilde{X}W$), where matrix $X_{T\times N}$ stacks N asset returns, and matrix $\Pi_{T\times K}$ stacks K factors, into their columns. Matrix $W_{N\times K}$ contains the weights of factors on assets in the model. Inversely, the loadings of assets on given factors can be estimated by a least-squares (linear regression) procedure $X_{tn} = \alpha_n + \sum_{k=1}^K \Pi_{tk}\beta_{kn} + \varepsilon_{tn}$, $n \in \{1,\ldots,N\}$, or in matrix form (stacking asset returns into columns),

$$X_{T\times N} = \mathbb{1}_{T\times 1}\alpha_{1\times N} + \Pi_{T\times K}\beta_{K\times N} + \varepsilon_{T\times N}. \tag{29}$$

Since G_K prices traded factors by construction, the model can also be characterized by a market-based SDF $M_{K\|}$ (with demeaned version $\widetilde{M}_{K\|}$) that is linear in the traded factors (i.e., the Hansen-

Jagannathan SDF projector),
$\widetilde{M}_{K\parallel} = -\widetilde{\Pi}\left[\frac{1}{T}\widetilde{\Pi}'\widetilde{\Pi}\right]^{-1}\mu'_{\Pi}$, where $1 \times K$ vector μ_{Π} contains the means (or risk premia) of K factors.[25]

To assess a pricing factor model, we employ four pricing measures that quantify various pricing performance aspects of the model and are standard to the literature. The four measures are (i) the maximum attainable Sharpe ratio, (ii) root-mean-square pricing error, (iii) Gibbons-Ross-Shanken (GRS) test statistic, and (iv) percentage of unexplained (idiosyncratic) return variation.

Maximum attainable Sharpe ratio (max SR): The max SR of the model G_K, denoted hereafter as $maxSR(G_K)$, is the highest possible SR that a portfolio spanned by its K factors can attain. Since G_K prices traded factors by construction, this max SR also equals the volatility of the market-based SDF $M_{K\parallel}$ defined above, $maxSR(G_K) = \sqrt{Var\left[M_{K\parallel}\right]}$. In the premise where factors $\{\Pi_n\}_{n=1}^{K}$ are mutually uncorrelated (or factors are re-orthogonalized) such as PCA, the max (squared) SR equals the sum of squared SR of factors

$$[maxSR(G_K)]^2 = \sum_{n=1}^{K}\left(SR\left[\Pi_n\right]\right)^2, \qquad \text{with} \quad \Pi_{kt} = \sum_{n=1}^{N} X_{nt}W_{nk}. \tag{30}$$

When all factors are retained ($K = N$), Proposition 1 implies that $maxSR(G)$ is invariant to the choice of return basis (and hence, also invariant to the set of factors resulted from that basis). When only a limited number of factors is retained ($K < N$), both the choices of asset return basis and retained factors matter for the $maxSR(G_K)$. That is, a need for the dimensionality reduction also motivates a performance comparison between different pricing models. Generally, a pricing model of higher max SR is desirable because it indicates that important risks (those having higher prices) are priced by the model.

Root-mean-square pricing error (RMSE): The pricing errors of assets in model G_K are the differences between assets' mean excess returns and assets' risk premia (priced by factors in the model). Since the risk premia are quantified by the covariance between asset returns and model's SDF, the pricing error α_n for a particular test asset X_n (or $1 \times N$ vector α for N asset tests in $\{X_n\}_{n=1}^{N}$) and their RMSE (denoted as $RMSE_\alpha$) in model G_K are respectively

$$\alpha_n = \mu_{xn} + \widetilde{M}'_{K\parallel}\widetilde{X_n} = \mu_{xn} - \mu'_{\Pi}\widetilde{\Pi}\left[\frac{1}{T}\widetilde{\Pi}'\widetilde{\Pi}\right]^{-1}\widetilde{X}_n, \qquad \alpha = \mu_x - \mu'_{\Pi}\widetilde{\Pi}\left[\frac{1}{T}\widetilde{\Pi}'\widetilde{\Pi}\right]^{-1}\widetilde{X}$$

[25]Recall that asset returns are excess returns. As a result, the means of traded factors that are linear in asset returns are also excess returns (risk premia) of the factors.

$$RMSE_{\alpha} = \frac{1}{N}\left[\sum_{n=1}^{N}\alpha_n^2\right]^{\frac{1}{2}} = \frac{1}{N}\sqrt{\alpha\alpha'}. \tag{31}$$

Evidently, a pricing model of lower $RMSE$ is desirable because it indicates a higher pricing ability of the model's factors. Note that $RMSE$ is amenable to the choice of return basis. Given an asset return space and the same set of factors $\{\Pi_n\}_{n=1}^{K}$, $RMSE_{\alpha}$ varies with a specific choice of test assets (spanning the given return space). Specifically, maintaining the same pricing model $\{\Pi_n\}_{n=1}^{K}$ while transforming the test assets by an uniform leverage operation on their excess returns $Y_n = \kappa X_n$, $n \in \{1, \ldots, N\}$ (2) scales linearly (hence arbitrarily, via the choice of κ), $RMSE_{\alpha}(Y) = \kappa RMSE_{\alpha}(X)$. With regard to this amenability, our empirical analysis employ both original and transformed returns as test assets for robustness. Similar to $maxSR(G_K)$, $RMSE_{\alpha}$ are computed in sample as well as out of sample, employing factors' in-sample weights but respectively in-sample and OOS asset returns.

GRS test statistic: Another pricing measure is the GRS test statistic, which generalizes the mean-variance efficiency test for the one-factor CAPM to a K-linear factor model employing N test assets (29). The GRS null hypothesis is that all pricing errors α's in (29) are jointly zero, or K factors suffice to capture all systematic risks impacting N test assets. Under the assumption that disturbances are jointly normally and independently distributed with zero means and $N \times N$ invertible covariance matrix Σ_{ε}, then the GRS test statistic defined as

$$\text{GRS} = \frac{T}{N} \times \frac{T-N-K}{T-K-1} \times \frac{\widehat{\alpha}\widehat{\Sigma}_{\varepsilon}^{-1}\widehat{\alpha}'}{1+\mu_{\Pi}'\widehat{\Sigma}_{\Pi}^{-1}\mu_{\Pi}}, \tag{32}$$

follow a F-distribution with N and $T-N-K$ degrees of freedom (or, $F_{N,T-N-K}$) under the null hypothesis. In (32), $\widehat{\Sigma}_{\Pi} = \frac{1}{T}\widetilde{\Pi}'\widetilde{\Pi}$, and $\widehat{\Sigma}_{\varepsilon} = \frac{\widehat{\varepsilon}'\widehat{\varepsilon}}{T-K-1}$, and $\widehat{\varepsilon}$ and $\widehat{\alpha}$ are least squares estimates from (29). A statistically sufficiently high value of GRS test statistic (quantified by a small enough p-value) is desirable as it indicates a no rejection for the null hypothesis (under which pricing errors are zero in the K-factor model).

Percentage of unexplained return variation: Under model G_K, the covariation between test asset returns and K factors are the explained (systematic) part of asset return variation, the remaining is unexplained (idiosyncratic) part. Given the representation (29) of the linear factor model, the unexplained part, denoted as σ_e, is computed as complementary component of the explained part (all in %) as follows

$$\sigma_e = 1 - \frac{\sum_{n=1}^{N}\sum_{k=1}^{K} Covar\left(X_n, \Pi_k \widehat{\beta}_{kn}\right)}{\sum_{n=1}^{N} Var\left(X_n\right)} = 1 - \frac{Trace\left[\widetilde{X}'\widetilde{\Pi}\widehat{\beta}\right]}{Trace\left[\widetilde{X}'\widetilde{X}\right]}, \tag{33}$$

where $\widehat{\beta}$'s are least squares estimates for the asset loadings on factors (29). When factors $\{\Pi_n\}_{n=1}^{K}$ are mutually uncorrelated (or factors are re-orthogonalized) such as PCs, the % of unexplained return variation becomes $\sigma_e = 1 - \frac{\sum_{k=1}^{K}\lambda_k}{\sum_{k=n}^{N}\lambda_N}$, where $\lambda_n = Var\left(\Pi_n\right)$, $n \in \{1, \ldots, N\}$. Evidently, a pricing model of lower σ_e is desirable (i.e., a larger part of test asset returns' variations can be explained by their exposures to K retained factors). However, the % of unexplained return variation σ_e does not explicitly account for pricing errors (i.e., first moment of returns) and is amenable to the choice of return basis and leverage operations on factors ($\Pi_n(\kappa_n) = \kappa_n\Pi_n$, $n \in \{1, \ldots, N\}$, i.e., $\Pi_n(\kappa_n)$ and Π_n represent the same risk. Accordingly, our empirical analysis examines moment-fusing constructions with return bases informed by first moment of returns, and given the amenability of σ_e and RMSE in this process, focuses on $maxSR$ as a primary pricing measure.

A.3 Equity Characteristics

Our equity data follows Lettau and Pelger (2020)'s employment of 74 and 370 extreme-decile anomaly portfolios (LP74 and LP370), which concern 37 underlying characteristics. They are: (1) Industry Relative Reversals, (2) Industry Momentum-Reversals, (3) Industry Relative Reversals, (4) Seasonality, (5) Value-Profitability, (6) 12-Month Momentum, (7) Value-Momentum-Profitability, (8) Investments Scaled by Assets, (9) Composite Issuance, (10) Investment Growth, (11) Sales/Price, (12) Earnings/Price, (13) Net Operating Assets, (14) Accrual, (15) Value (Annual), (16) Gross Profitability, (17) Asset Turnover, (18) Value-Momentum, (19) Cash Flows/Price, (20) Momentum-Reversals, (21) Asset Growth, (22) Long-Run Reversals, (23) Industry Momentum, (24) Idiosyncratic Volatility, (25) Value (Monthly), (26) Short-Term Reversals, (27) Size, (28) 6-Month Momentum, (29) Leverage, (30) Return on Assets, (31) Dividend/Price Ratio, (32) Investment/Capital, (33) Return on Book Equity, (34) Sales Growth, (35) Gross Margins, (36) Share Volume, (37) Price.

Among 37 characteristics listed above, 14 are constructed from price measures, including (1), (2), (3), (4), (6), (7), (18), (20), (22), (23), (26), (27), (28), (37). Those price measures are closely related to the momentum and reversal characteristics. Lettau and Pelger (2020) document that the portfolio returns constructed by those measures, especially when they are related to reversals, load heavily on the 5th RP-PCA factor that has a high Sharpe ratio of monthly values averaging around 0.46.

Other 21 characteristics among the listed are constructed from accounting measures, including (5), (8), (9), (10), (11), (12), (13), (14), (15), (16), (17), (19), (21), (25), (29), (30), (31), (32), (33), (34), (35). Valuation and profitability measures are the main components of the accounting characteristics in the sample and concern 9 out of those 21 characteristics, including (5), (11), (12), (15), (16), (25), (31), (33), (35). Lettau and Pelger (2020) document that the portfolio returns constructed by the valuation characteristics load heavily on the 3rd RP-PCA factor that has a modest Sharpe ratio of monthly values averaging around 0.06. Portfolio returns associated with other characteristics do not appear to load heavily on the 3th RP-PCA factor. Investment-related measures concern 4 of the 21 characteristics listed above, including (8), (10), (21), (32). Portfolio returns associated with these 4 characteristics do not load heavily on RP-PCA factors that have high Sharpe ratios Lettau and Pelger (2020). The trading liquidity concerns the other 2 characteristics, namely, (24) and (36).