

Moment Fusing: An Informed Construction of Pricing Factors ^{*}

Chun-Wei Lin[†]

Ngoc-Khanh Tran[‡]

Guofu Zhou[§]

October 18, 2026

Abstract

Standard PCA methods identify common risk factors while omitting information about expected returns. We propose a general, flexible and intuitive 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 scale with the Sharpe ratios of the original assets. We then perform a factor analysis on the transformed returns, obtaining factors that explain the cross section of Sharpe ratios of assets and price original asset returns. Empirically, we find that the moment-fusing factors significantly outperform leading benchmark models in pricing both FX and equity portfolios out-of-sample.

JEL-Classification: G10, G12, G15, F31.

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

^{*}We are very grateful to Torben Andersen, Tyler Beason, Hank Bessembinder, Jal Bhachech, Ric Colacito, Tuan Dinh, Roger Edelen, Jingzhi Huang, Chris Hrdlicka, Greg Kadlec, Raymond Kan, Raman Kumar, Anh Le, Greg Leonard, Alex Pecora, Ivan Shaliastovich, Andreas Stathopoulos, Allan Timmermann, Andrea Vedolin, Xiaolu Wang, Shixiang Xia, Pengfei Ye, an anonymous reviewer, and seminar participants at Virginia Tech, Washington University in St. Louis, 2026 China International Conference in Finance and 2026 Econometric Society North American Summer Meeting 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: k4956521@gmail.com.

[‡]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 conceptually elucidates this risk-return tradeoff ([Markowitz \(1952\)](#)), 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 possible 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. Intuitively, when SRs of original returns are large, their employment creates proportionally large volatilities for the transformed returns in the fusing stage. Large transformed volatilities then contribute strongly to the principal factors constructed from transformed returns in the factor analysis stage. As a result, the constructed principal factors are strongly influenced by the original risks of significant prices (i.e., large original SRs). Such constructed principal factors are uncorrelated, and delineated by their ability to explain 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 original 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 \(2020\)](#). 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. Conceptually, we consider 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. Empirically, we examine and demonstrate the out-of-sample pricing performance of moment-fusing factors disciplined by pricing restrictions in both foreign exchange (FX) and equity markets.

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 explain and reflect commensurate shares of SRs of the original returns. The ISR-PCA starts with transforming 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 explain and reflect commensurate shares of 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.

The moment-fusing approach is motivated and disciplined by several analytical properties. First, the moment-fusing constructions focus on delivering factors that explain SRs of the original assets. This focus is substantiated by a rotational invariance of PCs' Sharpe ratios (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), implying an identical sum of squares of principal factor prices in every return basis. Therefore, when only a limit-

¹A 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.

ed number of factors is retained (due to large data and dimensionality reduction requirement), the moment fusing prioritizes factors that explain large shares of SRs. The omitted (less prioritized) factors necessarily explain little SRs thanks to the SR rotational invariance mentioned earlier.

Second, moment fusing is general and flexible because the transformation underlying its first (fusing) stage can replicate any covariance structure of traded returns (Proposition 2).² We illustrate this flexibility by fusing (i.e., absorbing) a set of original pricing restrictions into the construction of transformed returns. Furthermore, not only different statistical moments (e.g., mean, skewness, kurtosis), but also various economic and pricing indicators (e.g., macro and firm characteristics, and risk loading β 's and mispricing α 's) of the original returns can be fused into the transformed returns.

Third, it is important to recognize that the mapping between moment-fusing factors and their factor prices is intuitive but not exact. As the moment-fusing approach is designed to enhance and identify the contribution of large original SRs, it has a limitation when the original SR distribution is not informative (e.g., original asset SRs are highly similar or clustered around smaller values). Given such a distribution, as the moment-fusing factors emphasize the explanation of SR clusters, a wedge between explaining a larger share of original SRs and having higher SRs can arise. To address this caveat and confront data, we incorporate and dynamically train pricing restrictions to improve the informativeness of the original SR distribution in an extended moment-fusing approach.

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.³ The analysis of the FX market employs the most actively traded currencies of 11 advanced economies and presents robust results for 4 different representative numeraire currencies (namely, the currency denominations of the US, Japan, Australia, and UK). The empirical analysis using FX data offers two advantages. First, the observable forward exchange rate discounts (or, the interest rate differentials) present a known proxy for

²Specifically, suppose that the asset return space has N dimensions. Then, given any $N \times N$ full-rank matrix Σ_y , we can always construct transformed asset returns, whose covariance matrix equals Σ_y .

³Four pricing measures are: the maximum attainable Sharpe ratio, root-mean-square pricing error, Gibbons-Ross-Shanken (GRS) test statistic, and share of unexplained (idiosyncratic) return variation.

the first moment of currency returns as documented in the currency carry trade literature. Second, the change of the numeraire currency adds flexibility and robustness to the moment-fusing constructions by preserving the volatility (risk) space but altering the pricing (risk premia) perspective. We find that the moment-fusing and extended pricing models outperform consistently in various numeraire currencies. Specifically, the highest annualized OOS maximum attainable SRs obtained by an extended moment-fusing construction with two retained factors in 4 numeraire currencies are 0.28 (for USD), 0.72 (for JPY), 1.13 (for AUD), and 0.65 (for GBP), extending [Lustig et al. \(2011\)](#)'s seminal two-factor FX pricing model of respective SRs 0.09 (USD), 0.18 (JPY), 0.01 (AUD), and 0.11 (GBP).

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 first moment (or its proxy) of these original portfolios and pricing restrictions to inform the transformed returns for the construction of moment-fusing and extended pricing models. The benchmark model for the equity market employs the pricing restrictions (also dynamically trained, but without moment fusing). We retain either three or five factors in every model. The extended moment-fusing and related factors outperform across different numbers of retained factors and data sets. Specifically, the highest monthly OOS maximum attainable SRs obtained by an extended moment-fusing construction in 74-, 370-, and 25-portfolio data set are respectively 0.18, 0.21, and 0.42 (for three retained factors) and 0.55, 0.37, and 0.51 (for five retained factors). The corresponding values for the benchmark model are 0.10, 0.19, and 0.25 (for three retained factors) and 0.45, 0.37, and 0.34 (for five retained factors), extending the moment-fusing approach with [Lettau and Pelger \(2020\)](#)'s prominent pricing restrictions to improve pricing performance.

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) 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 benefits from employing pricing restrictions to strengthen pricing performance OOS. In the difference with the factor construction literature, moment-fusing approach proposes a direct input of prices of risks, and possibly any other moments and characteristics of asset returns, into the construction. Our paper also follows [Maurer et al. \(2022\)](#) and [Maurer et al. \(2023\)](#), who document that factors levered by risk prices in FX markets offer improvements in both pricing performance and predictability. Motivated by the findings of [Bessembinder et al. \(2025\)](#) that same factor construction performs differently in different data sets and different factor constructions perform differently in same data set, we examine moment-fusing factors (and alternative factor constructions) in various FX and equity data sets. While our empirical analysis indeed finds that pricing models perform differently in FX and equity markets, the moment-fusing and related constructions consistently outperform leading benchmark models. In particular, the moment-fusing approach extends [Lustig et al. \(2011\)](#)'s prominent two-factor pricing model of the FX market to other denomination currencies (beyond the USD), following [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. [Bessembinder et al. \(2022\)](#) document a significant time variation in the number of pricing-relevant PCs in the equity market. The moment-fusing implementation accounts for the time-varying fac-

tor structure by dynamically estimating factor loadings and other moment-fusing parameters in expansive windows. Recent aspects of assessing the performance of pricing models are discussed in [Barillas and Shanken \(2017\)](#) concerning pricing errors and [Kan et al. \(2024\)](#) concerning attainable Sharpe ratios. 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 pricing errors and OOS 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 simple illustrations. Section 3 quantifies specific moment-fusing constructions and extends these constructions by incorporating the pricing restrictions. Section 4 discusses conceptual properties of the moment-fusing approach. Section 5 presents the empirical analysis of the moment-fusing constructions using FX data and Section 6 using equity data. Section 7 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 to traded asset returns. Section 2.1 introduces a market setup of generic traded asset returns and notations. Section 2.2 presents simple illustrations motivating the subsequent moment-fusing approach.

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 $t \in \{0, \dots, T\}$, 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 approach, risky assets are currency strategies (Section 5) and equity portfolios (Section 6). For the holding period from t to $t + 1$, let

the bond's return and risky assets' excess returns respectively be

$$(1) \quad B_{0t+1} = r_t, \quad X_{nt+1} = \mu_{xn} + \sigma'_{xn} \varepsilon_{xt+1}, \quad n \in \{1, \dots, N\}, \quad t \in \{0, \dots, T-1\},$$

where r_t 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 expositional simplicities, we omit their time variation and time index.⁴ Throughout, $\tilde{A}$ 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 enables constructions of new asset bases that retain certain characteristics of the original one without altering the return space or the underlying model. Below, we generically refer an equivalent asset basis (and its returns) $\{B_0, Y_n\}_{n=1}^N$ as a transformed asset basis (and transformed asset returns).

One simple set of transformed assets concerns a pure leverage operation (i.e., scaling), in which each transformed asset Y_n is a portfolio of only the respective original risky asset X_n (with weight κ_n) and the risk-free bond B_0 (with weight $1 - \kappa_n$),

$$(2) \quad Y_{nt+1}^{\text{Full}}(\kappa_n) = \kappa_n X_{nt+1}^{\text{Full}} + (1 - \kappa_n) B_{0t+1} \implies \underbrace{Y_{nt+1}^{\text{Full}}(\kappa_n) - B_{0t+1}}_{Y_{nt+1}(\kappa_n)} = \kappa_n \underbrace{(X_{nt+1}^{\text{Full}} - B_{0t+1})}_{X_{nt+1}},$$

where X_{nt+1}^{Full} and Y_{nt+1}^{Full} respectively are the full (non-excess) return versions of the original and transformed excess returns X_{nt+1} and Y_{nt+1} in (1). Evidently, the linearly scaled excess return $\kappa_n X_{nt+1}$ remains a traded excess return for all real parameters $\kappa_n \neq 0$. The substitution of the

⁴The moment-fusing approach 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. Note that the transformation of asset returns (i.e., the first stage of the moment fusing, (4) below) $S(m)$ passively fuses the time variation in return moments $S(m_t)$, but can also be actively constructed to be time-varying, $S_t(m_t)$. E.g., [Maurer et al. \(2022\)](#) document that FX factors, which are levered by time-varying risk prices, offer superior pricing performance.

original excess returns (1) into the above pure leverage operation then implies the return moments $\{\mu_{yn}, \sigma_{yn}\}$ of the transformed assets

$$(3) \quad Y_{nt+1}(\kappa_n) = \kappa_n X_{nt+1} = \underbrace{\kappa_n \mu_{xn}}_{=\mu_{yn}} + \underbrace{\kappa_n \sigma'_{xn} \varepsilon_{xt+1}}_{=\sigma'_{yn} \varepsilon_{yt+1}}, \quad n \in \{1, \dots, N\}, \kappa_n \in \mathbb{R}.$$

Note that $Y_{nt+1}(\kappa_n)$ 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 $\kappa_n \neq 0$. More generally, observe that any 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. This flexibility is the basis for our moment-fusing approach, in which return transformations help to identify important pricing factors. As every return transformation is composed of rotations and (pure-leverage) scaling, among which only the scaling is able to alter the associated principal factors (discussed further in Equation (18), Section 4.1 below). Therefore, while appearing simple, the pure leverage operations (3) are rich to capture the essence of the general moment-fusing approach in identifying and delivering capable transformed pricing factors.

2.2 Moment Fusing and Preliminary Illustrations

Moment fusing is a general, systematic and flexible approach 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 m_x of original returns. In the second stage, a covariance-based factor analysis is performed on the transformed returns to identify factors $\Pi_y(m_x)$ that reflect

original moments and price original returns.

Moment Fusing:

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

where $\Sigma_x = \frac{1}{T} \widetilde{X}' \widetilde{X}$ is the covariance matrix of the original returns and throughout the notation $\text{Diag} [\lambda]$ denotes the diagonal matrix with diagonal entries being the components of vector λ . 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 approach, the first stage preserves the return space (hence, the underlying model), the second stage retains and utilizes advantages of the powerful factor analysis methodology. The fact that factors constructed from an equivalent basis of transformed returns offer a differentiated (possibly improved) pricing performance compared to those constructed from the original basis of the same return space supports and strengthens [Bessembinder et al. \(2025\)](#)'s empirical findings that pricing performance varies with both factor construction and data set.

Below, we present three illustrations for the moment fusing. They are based on known and simple asset transformations but do not yet possess pricing powers. These preliminary illustrations provide motivations for more elaborate moment-fusing constructions of desirable pricing properties 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 a unit variance. Employ-

ing $\kappa_n = \frac{1}{|\sigma_{xn}|}$ in the transformation (3) produces

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

where $\hat{\sigma}_{xn}$ is a unit vector ($|\hat{\sigma}_{xn}| = 1$). Evidently, $Corr_t[Y_{nt+1}(\kappa_n), Y_{kt+1}(\kappa_k)] = Cov_t[X_{nt+1}(\kappa_n), X_{kt+1}(\kappa_k)]$, for all $n, k \in \{1, \dots, N\}$. 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 $\kappa_n = \frac{1}{|\sigma_{xn}|}$, $n \in \{1, \dots, N\}$, 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 $\tilde{X}'\tilde{X}$ quantifies the return comovements around the means of respective asset returns and employs exclusively the demeaned asset returns $\{\tilde{X}_n\}_{n=1}^N$. 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 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 moment 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.⁵ 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 remains the same, resulting in the same PCs for the two (original and transformed) return bases X and Y . As a 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 present such moment-fusing constructions next.

⁵ The covariance matrices of transformed and original returns are related as, $\Sigma_y = \frac{1}{T}\widetilde{Y}'\widetilde{Y} = \frac{1}{T}Q'\widetilde{X}'\widetilde{X}Q = Q'\Sigma_xQ$. If Σ_x is diagonalized by W_x , or $W_x'\Sigma_xW_x = \text{Diag}(\lambda_x)$, then Σ_y is diagonalized by $W_y = Q'W_x$. Indeed, $(W_y')(\Sigma_y)(W_y) = (W_x'Q)(Q'\Sigma_xQ)(Q'W_x) = W_x'\Sigma_xW_x = \text{Diag}(\lambda_x)$. 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: Specific Constructions

This section introduces specific moment-fusing constructions by transforming the return volatilities into SRs (Section 3.1), standardizing the mean returns (Section 3.2), and extending 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. The construction equalizes the second moment of the transformed asset returns 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 (3) and produces the transformed returns

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

where $SR[X_n] = \frac{\mu_{xn}}{|\sigma_{xn}|}$ is the Sharpe ratio of n th 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 $S(m_x) = \text{Diag}\left(\frac{(\mu_x)^{\frac{1}{2}}}{|\sigma_x|^{\frac{3}{2}}}\right)$, and the transformed covariance matrix Σ_y is related to the original counterpart Σ_x as

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

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

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. Though transformations (6) are simple (pure-leverage) scaling that concerns only a pair of asset returns $\{X_n, Y_n\}$ at a time, they suffice to generate a rich set of moment-fusing pricing factors (as quantified by Equation (18) below).

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),

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

where $\{\lambda_{yn}\}$ are eigenvalues of the 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, the design of SRM transformation aims to enhance and identify the contribution of large original SRs. Hence, it has a limitation when original SRs do not exhibit an informative (i.e., heterogeneous) pattern. Such an “uninformative” premise takes place when original SRs

are similar. In this premise, the transformed returns $\{Y_{nt+1}\}$ (6) have similar (indistinctive) variances. As a result, while PCs obtained from $\{Y_{nt+1}\}$ are prioritized to explain the covariation of original SRs, these PCs do not necessarily have large SRs if these original SRs are clustered around smaller values. That is, $\{\lambda_{yn}\}$ do not always sort monotonically with $\{SR[\Pi_{yn}]\}$. Empirically, we incorporate pricing restrictions into moment-fusing constructions to mitigate this caveat (discussed further in Section 4.1). It can also be addressed by scaling (pure leverage) operations, which change the volatilities $\{\lambda_{yn}\}$ but leaves factor prices $\{SR[\Pi_{yn}]\}$ intact (as in Section 4.2).

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

$$(9) \quad Y_{nt+1}(\kappa_n) = \kappa_n X_{nt+1} = \mu + \mu \frac{1}{SR[X_n]} \hat{\sigma}'_{xn} \varepsilon_{xt+1},$$

$$\implies \quad \mu_{yn} = \mu, \quad \sigma_{yn} = \mu \frac{|\sigma_{xn}|}{\mu_{xn}} = \mu \frac{1}{SR[X_n]}, \quad n \in \{1, \dots, N\},$$

where μ is a notional constant parameter.⁶ In this construction, up to the non-material multi-

⁶The specific value of μ has no material effect on the moment-fusing procedure or the resulting factors as it is a

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

$$\frac{1}{T} \widetilde{Y}' \widetilde{Y} = \begin{bmatrix} \frac{1}{\mu_{x1}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\mu_{xN}} \end{bmatrix} \frac{\widetilde{X}' \widetilde{X}}{T} \begin{bmatrix} \frac{1}{\mu_{x1}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\mu_{xN}} \end{bmatrix},$$

$$(10) \quad \text{or,} \quad \Sigma_y = \underbrace{\text{Diag}\left[\frac{1}{\mu_x}\right] \Sigma_x \text{Diag}\left[\frac{1}{\mu_x}\right]}_{\equiv ISRM(X)},$$

Hereafter, we refer to $\Sigma_y = ISRM(X) \equiv \text{Diag}\left[\frac{1}{\mu_x}\right] \Sigma_x \text{Diag}\left[\frac{1}{\mu_x}\right]$ 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, $ISRM(X)$ (10) is a moment-fusing construction as the original means $\{\mu_{xn}\}_{n=1}^N$ are incorporated into the transformed volatilities $\{\sigma_{yn}\}_{n=1}^N$ (9). Eigenvalues and eigenvectors of Σ_y then are also fused with original mean returns. Note that the moment-fusing pattern of $ISRM(X)$ (10) differs from the previous one of $SRM(X)$ (7). This is evidenced in the fact that while being related, $ISRM(X)$ strictly differs from $[SRM(X)]^{-2}$.⁷

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 (9), 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 name inverse SR matrix). Building on this intuition and applying PCA on the inverse SR ma-

common multiplicative factor for all returns $\{Y_n\}_1^N$.

⁷Similar to the SR-PCA transformation (6), the pure-leverage ISR-PCA transformation suffices to generate a rich set of moment-fusing pricing factors as explained in Equation (18), Section 4.1 below.

trix $ISRM(X)$ (10), the moment-fusing construction prioritizes less volatile factors among PCs $\{\Pi_{yn}\}$ associated with the transformed return basis $\{Y_n\}$ (9),

$$(11) \quad \textbf{Inverse Sharpe Ratio-PCA (ISR-PCA):} \quad \lambda_{yn} < \lambda_{yk} \longrightarrow \Pi_{yn} \succ \Pi_{yk},$$

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 5 and 6) retains only eigenvalues $\{\lambda_{ny}\}$ above a threshold sufficiently different from zero, before applying the ISR-PCA prioritization (11). 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, $\tilde{Z} \equiv \tilde{X} [\Sigma_x]^{-1} \text{Diag}[\mu_x]$. 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.⁸ Finally, we recall that the transformed mean returns (9) 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 Extended 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

⁸Specifically, because matrix Σ_z and its inverse $\Sigma_z^{-1} = \Sigma_y$ are diagonalized by the same rotation matrix W_y , we have $\text{Diag}(\lambda_z) = W_y' \Sigma_z W_y = W_y' \Sigma_y^{-1} W_y = \text{Diag}(\frac{1}{\lambda_y})$. 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 moment (and SRs) of the original return X .

factors. Building upon RP-PCA framework, this section extends the moment-fusing approach to incorporate the pricing restrictions into ISRM-PCA and SRM-PCA constructions. An in-depth discussion of the relation between RP-PCA and moment-fusing approaches and an empirical assessment of pricing models that combine these approaches are relegated respectively to Section 4 and Sections 5 and 6.

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 a weighted objective function of unexplained covariation (concerning the second moment) and pricing errors (i.e., pricing restrictions, concerning the first moment) of asset returns

$$(12) \quad \min_{\{\Pi^{RP}, W^{RP}\}} \text{Trace} \left\{ \frac{1}{T} [\tilde{X} - \tilde{\Pi}^{RP} W^{RP'}]' [\tilde{X} - \tilde{\Pi}^{RP} W^{RP'}] + (1 + \gamma) [\bar{X} - \bar{\Pi}^{RP} W^{RP'}]' [\bar{X} - \bar{\Pi}^{RP} W^{RP'}] \right\},$$

where $\tilde{X}_{T \times N}$ and $\bar{X}_{1 \times N}$ denote respectively matrices of (demeaned) asset returns and their means, $\tilde{\Pi}_{T \times N}^{RP}$, $\bar{\Pi}_{1 \times N}^{RP}$, and $W_{N \times N}^{RP}$ 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.⁹ Lettau and Pelger (2020) show that RP-PCA factor construction is equivalent to diagonalizing a $N \times N$ risk-premium adjusted 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 $\Sigma_x^{RP}(\gamma)$

$$(13) \quad \Pi^{RP}(\gamma) = X W^{RP}(\gamma), \quad \text{with} \quad \begin{cases} \Sigma_x^{RP}(\gamma) \equiv \frac{1}{T} \tilde{X}' \tilde{X} + (1 + \gamma) \bar{X}' \bar{X}, \\ W^{RP'}(\gamma) \Sigma_x^{RP}(\gamma) W^{RP}(\gamma) = \text{Diag} [\lambda_n^{RP}(\gamma)], \end{cases}$$

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

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

⁹In comparison, the standard PCA identifies factors (or principal components) by minimizing only the unexplained covariation of asset returns, $\min_{\{\Pi^{RP}, W^{RP}\}} \text{Trace} \left([\tilde{X} - \tilde{\Pi}^{RP} W^{RP'}]' [\tilde{X} - \tilde{\Pi}^{RP} W^{RP'}] \right)$. As a result, RP-PCA coincide with PCA when $\gamma = -1$.

returns. Hereafter, we omit the explicit γ -dependence notation attached to RP-PCA's quantities when the omission does not create ambiguity.

Integrating pricing restrictions into moment fusing: We extend the moment-fusing approach to incorporate Lettau and Pelger (2020)'s pricing restrictions into the second (factor analysis) stage of the approach (4).¹⁰ Specifically, after constructing the transformed returns in the first stage, we optimize the weighted objective function (12) 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 employing a similar deduction from the optimization (12) to the diagonalization of an adjusted covariance matrix (13), the second stage of the moment-fusing approach (4) extended with pricing restrictions amounts to diagonalizing the adjusted matrix $\Sigma_y(m_x) = \frac{1}{T} \tilde{Y}' \tilde{Y} + (1 + \gamma) \bar{Y}' \bar{Y}$.

In summary, extending the moment fusing (4) with pricing restrictions (13), the pricing factors in the $T \times N$ matrix $\Pi_y(\gamma, S_m)$ are constructed as follows

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

$$(14) \quad \textbf{Extended moment fusing: } \lambda_{yn}(\gamma, m_x) > \lambda_{yk}(\gamma, m_x) \longrightarrow \Pi_{yn}(\gamma, m_x) \succ \Pi_{yk}(\gamma, m_x).$$

Notationally, the constructed factors $\Pi_y(\gamma, m_x)$ feature both moment fusing (m_x) and pricing restriction (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 em-

¹⁰As discussed further in Section 4.2, such an extension also helps to mitigate a possible wedge between factor prices and factor volatilities in moment fusing.

ploy the respective moment-fusing matrix $S(m_x) = \text{Diag} \left(\frac{(\mu_x)^{\frac{1}{2}}}{|\sigma_x|^{\frac{3}{2}}} \right)$ in the construction (14). When applying the extended moment-fusing construction to the ISR-PCA (10), we employ the respective $S(m_x) = \text{Diag} \left(\frac{1}{\mu_x} \right)$.¹¹ Empirically, to utilize the moment-fusing flexibility, we further combine the factors obtained from the extended moment-fusing constructions (14). The combined factors are either (i) equally weighted, or (ii) optimally weighted

$$(15) \quad \begin{cases} \text{Equally weighted:} & \Pi_{ew} = 0.5 \Pi_{SR-PCA}(\gamma_e, m_x) + 0.5 \Pi_{ISR-PCA}(\gamma_e, m_x), \\ \text{Optimally weighted:} & \Pi_{ow} = (1 - w^*) \Pi_{SR-PCA}(\gamma_o, m_x) + w^* \Pi_{ISR-PCA}(\gamma_o, m_x), \end{cases}$$

wherein the optimal weight w^* is dynamically determined using expansive windows. Sections 5 and 6 below present implementation details and empirical results for FX and equity markets. Section 4 presents further conceptual properties of the moment-fusing approach.

4 Qualitative Analysis

This section presents qualitative aspects of the moment-fusing and related approaches. Section 4.1 details several important properties and a caveat of the moment-fusing approach. Section 4.2 relates moment fusing to RP-PCA framework and motivates a combination of the two approaches.

4.1 Moment Fusing: Properties

We illustrate moment-fusing properties by examining (i) the ability in explaining Sharpe ratios, and (ii) the generality and flexibility, of the moment-fusing factors.

Sharpe ratio-based prioritization: We observe that aiming for a higher SR explanation motivates the moment-fusing constructions of pricing factors. As the second stage of the moment-fusing process is a PCA applied to the transformed returns $\{Y_n\}$, the resulting factors $\{\Pi_{yn}\}$ max-

¹¹ Furthermore, because the transformed volatilities are inverse original SRs in the ISR-PCA model, the objective of finding common risk factors that reflect large original SRs is effectuated by changing the sign of the centered covariance term in the model's adjusted transformed covariance matrix. That is, in the extended ISR-PCA model, we flip the sign of the centered covariance term $\frac{1}{T} \tilde{Y}' \tilde{Y}$ in transformed covariance matrix $\frac{1}{T} \tilde{Y}' \tilde{Y} + (1 + \gamma) \bar{Y}' \bar{Y}$ while also using the standard eigenvalue ranking (prioritizing larger eigenvalues). This simultaneous adjustments reconcile the ISR-PCA reverse eigenvalue ranking (11) with the standard eigenvalue ranking in (14).

imize the explained (equivalently, minimize the unexplained) covariation of the transformed returns,

$$(16) \quad \min_{\{\Pi_y, W_y\}} \text{Trace} \frac{1}{T} [\tilde{Y} - \tilde{\Pi}_y W_y']' [\tilde{Y} - \tilde{\Pi}_y W_y'] .$$

By construction of the first stage, the transformed returns' volatilities (embedded in $\tilde{Y}$) are various measures of original SRs (see (6) and (9)). As a result, it is important to observe that the moment-fusing objective and resulting factors maximize the explained the covariation of SRs of original returns (16).¹²

Consider the SR-PCA construction for an illustration. Given the transformed return volatilities $\sigma_{yn} = (SR[X_n])^{1/2}$, $\forall n \in \{1, \dots, N\}$ (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 in PCA, this total variance also equal the sum of all principal eigenvalues,¹³ 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 $\{\Pi_{yi}\}_{i=1}^K$, the share of the explained transformed return covariation by these K PCs is

$$(17) \quad \frac{\sum_{n=1}^N \text{Cov}(Y_n, \sum_{i=1}^K \Pi_{yi} W_{yin}')}{\sum_{n=1}^N \text{Var}(Y_n)} = \frac{\sum_{i=1}^K \lambda_{yi}}{\sum_{n=1}^N \lambda_{yn}} = \frac{\sum_{i=1}^K \text{Var}(\Pi_{yi})}{\sum_{n=1}^N SR[X_n]} .$$

That is, a factor Π_{yi} associated with a larger value λ_{yi} in the SR-PCA ranking (8) is responsible for a larger explained share of the total original assets' SRs.¹⁴ The following observation substantiates this SR-based prioritization of moment-fusing factors further (Appendix B.1 presents a derivation of this result).

Proposition 1 (Sharpe Ratio Rotational Invariance). *The factor prices (i.e., SRs) of PCs constructed from any two equivalent asset bases $\{X_n\}_{n=1}^N$ and $\{Y_n\}_{n=1}^N$ of a return space are always related by a*

¹²Since the second stage is a PCA, the limiting distributions and efficiency of the moment-fusing factor estimates in terms of the transformed return Y 's distribution assumptions follow Bai (2003)'s standard treatment. When the moment-fusing approach features a predeterminedistic and well-behaved transformation matrix $S(m)$, these limiting distributions of factor estimates can be translated in terms of the original return X 's distribution assumptions.

¹³This is a standard PCA property. As returns $\{Y_n\}_{n=1}^N$ and PCs $\{\Pi_{yn}\}_{n=1}^N$ are related by a rotation matrix, $\Pi_y = Y W_y$ or $Y = \Pi_y W_y'$, their total variances are the same, $\sum_{n=1}^N \sigma_{yn}^2 = \text{Trace}\{\tilde{Y}'\tilde{Y}\} = \text{Trace}\{W_y \tilde{\Pi}_y' \tilde{\Pi}_y W_y'\} = \text{Trace}\{\tilde{\Pi}_y' \tilde{\Pi}_y W_y' W_y\} = \text{Trace}\{\tilde{\Pi}_y' \tilde{\Pi}_y\} = \text{Trace}\{\text{Diag} \lambda_y\} = \sum_{n=1}^N \lambda_{yn}$.

¹⁴The rationale for ISR-PCA factors is similar. By construction (9), these factors explain a share of total original inverse SRs. Since risks of little prices (i.e., noise risks of small SRs) in original asset returns dominate the total original inverse SRs, the reverse ranking (11) aims to disassociate the prioritized ISR-PCA factors from these noise risks, those are unimportant to the pricing of original returns.

rotation, $SR[\Pi_y] = Q' SR[\Pi_x]$, where Q is a $N \times N$ rotation matrix, and the $N \times 1$ vector $SR[\Pi_m]$ contains factor prices of N PCs in $\{\Pi_m\}$.

Recall from the leverage operation (3) that means and variances of principal factors can be linearly scaled by arbitrary factors. In contrast, Proposition 1 asserts that principal factors' prices are subject to a rotational invariance, hence cannot be arbitrarily scaled. As a result, when only a limited number of PC factors is retained (i.e., dimensionality reduction), this rotational invariance assures that the higher prices the retained factors have, the lower prices the omitted (not retained) factors have, regardless of specific asset return basis. This invariance motivates the prioritization and retention of moment-fusing factors that explain a larger share (17) of total SRs of asset returns.

Generality of moment-fusing constructions: Note that while the general moment fusing $Y = XS(m)$ (4) involves a generic full-rank matrix $S(m)$, SR-PCA (6) and ISR-PCA (9) constructions concern only a scaling of two assets $\{Y_n = X_n \kappa_n(m)\}$ at a time, separately for each $n \in \{1, \dots, N\}$. To see whether such SR-PCA and ISR-PCA constructions are too restrictive, or more importantly, how to discipline the generality of the moment-fusing transformation $S(m)$, we recall the following decomposition. Namely, every $N \times N$ full-rank matrix S can always be decomposed into a product of one scaling and two rotation matrices,

$$(18) \quad \underbrace{S}_{N \times N} = \underbrace{Q_L}_{N \times N} \underbrace{Diag(\lambda_s)}_{N \times N} \underbrace{Q_R}_{N \times N}, \quad \text{with} \quad \begin{cases} Q_L' Q_L = Q_L Q_L' = \mathbb{1}_{N \times N}, \\ Q_R' Q_R = Q_R Q_R' = \mathbb{1}_{N \times N}, \end{cases}$$

where the diagonal (scaling) matrix $Diag(\lambda_s)$ has non-zero elements only in its diagonal, and Q_L and Q_R are two orthogonal (rotation) matrices. The general decomposition (18) implies that any equivalent basis Y of a linear return space can be achieved by rotating, then scaling (stretching or shrinking) and rotating again, the original basis X . Two observations are in order. First, given a general asset return transformation S and its decomposition (18), the scaling component $Diag(\lambda_s)$ is essential to the construction and variation of the transformed principal factors. This is because rotation components $\{Q_L, Q_R\}$ leave PC factors unchanged (as observed in Section 2.2 and Footnote 5). Second and relatedly, because pure leverage operations (2) are scaling transformations, they suffice to capture the essence of transformed factors' construction. Hence, a

general moment-fusing transformation $S(m)$ can be disciplined and specified by focusing on its the scaling (or, pure leverage) component. These observations show that while SR-PCA and ISR-PCA are constructed using apparently only pure leverage operations (6) and (9), these constructions are general in terms of identifying model-fusing factors.

Moment-fusing designs: A central component of a moment-fusing process (4) is the transformation matrix $S(m)$, which encodes information about original return moments m (first stage) into the transformed return covariance matrix $\Sigma_y = S'(m)\Sigma_x S(m)$ for the identification of moment-fusing factors Π_y (second stage). While moment-fusing SR-PCA (6) and ISR-PCA (9) constructions start out with specific and explicit matrices $S(m)$ of desirable features, the design of a general moment-fusing construction can be more flexible. The following observation substantiates this flexibility in moment-fusing designs (Appendix B.1 presents a derivation of this result).

Proposition 2 (Moment-fusing Design). *Consider a set of N non-redundant (original) risky assets $\{X_n\}_{n=1}^N$. Given any $N \times N$ symmetric, real, and positive definite matrix Σ_y , there are always N (transformed) risky assets $\{Y_n\}_{n=1}^N$ that (i) span the same return space as the original assets, and (ii) have covariance matrix Σ_y . Specifically, the transformed and original asset returns are related by a full-rank matrix,*

$$(19) \quad Y = XS, \quad \text{with} \quad S = W_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} \left[\sqrt{\lambda_y} \right] W_y',$$

where Q is an arbitrary $N \times N$ rotation matrix, and W_x and W_y respectively diagonalize the original and transformed covariance matrices $W_x' \left(\frac{1}{T} \tilde{X}' \tilde{X} \right) W_x = \text{Diag} [\lambda_x]$, and $W_y' \Sigma_y W_y = \text{Diag} [\lambda_y]$.

Proposition 2 shows that return transformations are able to generate any regular covariance structure without changing the return space, and hence, without changing the underlying asset pricing model. This result gives rise to the flexibility for moment-fusing designs. That is, as the moment-fusing transformation S (19) is informed and guided by Σ_y , choosing a desirable transformed covariance structure Σ_y helps to design the corresponding desirable moment-fusing model. Note that the presence of a rotation matrix Q in (19) signifies that there are multiple transformed return solutions Y to a given covariance structure Σ_y . Below, we employ this flexibility to relate the RP-PCA framework and moment-fusing approach (Section 4.2,

Equation (21)).

Discussion: The moment-fusing constructions of pricing factors are motivated by an intuition that, all else equal, factors explaining a larger share of original asset SRs (as in the ratio (17) for SR-PCA factors) tend to represent original asset risks of higher SRs. However, it is important to recognize a caveat that this intuition may break down when the distribution of original SRs is uninformative. Specifically, when original SRs are either similar or clustered around smaller values, SR-PCA and ISR-PCA constructions do not necessarily deliver factors of highest factor prices. For such SR distributions, a wedge between explaining a larger share of SRs of asset returns and having higher SRs can arise. Three exemplifying settings help to discern the informativeness of the SR distribution. First, consider a set of many highly correlated asset returns that also have small SRs (the majority of returns) and a few uncorrelated asset returns that also have high SRs (the minority of returns). Top moment-fusing factors obtained from this original return set do not necessarily have high SRs because they tend to highly correlate with (i.e., explain) the majority of returns, which have small SRs in the setting. Second, when original SR distribution is more informative (e.g., spreads more broadly or clusters around larger values), top moment-fusing factors are influenced more by more volatile transformed returns, load more on original returns of higher SRs, and have higher SRs. Third, when asset returns are uncorrelated (or almost uncorrelated), top SR-PCA and ISR-PCA factors unambiguously have higher SRs.¹⁵ To mitigate the above-mentioned caveat, we incorporate pricing restrictions to address the informativeness of the original SR distribution in an extended moment-fusing approach.

4.2 Pricing Restrictions and Moment Fusing

While both [Lettau and Pelger \(2020\)](#)'s prominent RP-PCA framework and the moment-fusing approach incorporate information from the first moment of returns into factor constructions, the incorporation methods and resulting factors differ. It is instructive to examine these differences, before we extend the moment-fusing constructions by adopting the RP-PCA's pricing restrictions to confront data. Below, we discuss the objectives, relationships and scaling properties of

¹⁵When returns are uncorrelated, PCA leaves these returns intact (i.e., uncorrelated returns are PCs), and principal eigenvalues are asset return variances. Hence, principal eigenvalues are SRs for SR-PCA, and inverse SRs for ISR-PCA. As a result, SR-PCA and ISR-PCA rankings (8), (11) then unambiguously prioritize PCs of superior SRs.

RP-PCA and moment-fusing approaches.

RP-PCA and moment-fusing objectives: The difference in RP-PCA and moment-fusing objectives gives rise to differences in the resulting factors. RP-PCA factor constructions aim to minimize a weighted combination (12) of unexplained return covariation and pricing errors, and hence, concerns a linear combination $L(\sigma_x, \gamma\mu_x)$ of second and first moments of asset returns. Moment-fusing constructions aim to minimize the unexplained covariation of original SRs (16)-(17), and hence, concerns a non-linear combination (i.e., a ratio $\frac{\mu_x}{\sigma_x}$) of first and second moments of original asset returns. Given these perspectives, the RP-PCA objective can be elucidated by examining whether the RP-PCA prioritization delivers higher SR. To this end, we explicitly compute the prices of RP-PCA factors,¹⁶

$$(20) \quad SR(\Pi_n^{RP}) = \frac{\mu_{\Pi_n}^{RP}}{\sqrt{Var(\Pi_n^{RP})}} = \left[\frac{\lambda_n^{RP}}{(\mu_{\Pi_n}^{RP})^2} - (\gamma + 1) \right]^{-\frac{1}{2}},$$

where $\mu_{\Pi_n}^{RP} = \bar{X}W_n^{RP}$ is the mean return of the n -th RP-PCA factor, and W_n^{RP} is the n -th column of rotation matrix W^{RP} (13). Evidently, RP-PCA factor price $SR(\Pi_n^{RP})$ is inversely related to the ratio $\frac{\lambda_n^{RP}}{(\mu_{\Pi_n}^{RP})^2}$, 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 always have largest prices $SR(\Pi_n^{RP})$.¹⁷ The non-monotonic relationship between RP-PCA SRs and eigenvalues indicates improvements by combining RP-PCA and moment-fusing approaches.

Alternatively, we can gain insights into RP-PCA and moment-fusing objectives in special limits, in which simple relationships between respective covariance matrices exist.¹⁸ In particular, in the limit where risky assets have similar mean returns (i.e., controlling for mean returns and

¹⁶Recall that λ_n^{RP} is not the variance of n -th RP-PCA factor Π_n^{RP} . Instead, $Var(\Pi_n^{RP}) = E[(\Pi_n^{RP})^2] - E^2[\Pi_n^{RP}] = \frac{1}{T}\Pi_n^{RP'}\Pi_n^{RP} - (\mu_{\Pi_n}^{RP})^2 = W_n^{RP'}\frac{1}{T}X'XW_n^{RP} - (\mu_{\Pi_n}^{RP})^2$. To compute this variance, note from (13) that $X'X = \frac{1}{T}\tilde{X}'\tilde{X} + \bar{X}'\bar{X} = \Sigma_x^{RP} - \gamma\bar{X}'\bar{X}$, which then implies $W_n^{RP'}\frac{1}{T}X'XW_n^{RP} = \lambda_n^{RP} - \gamma(\mu_{\Pi_n}^{RP})^2$. As a result, the variance of RP-PCA factor is, $Var(\Pi_n^{RP}) = \lambda_n^{RP} - (\gamma + 1)(\mu_{\Pi_n}^{RP})^2$, which delivers the factor price in (20).

¹⁷Note that the RP-PCA minimization of pricing errors and unexplained covariation of asset returns (12) is equivalently characterized by the diagonalization (13), which is not equivalent to the determination of factor risk premia $\{\mu_{\Pi_n}^{RP}\}$. That is, leading RP-PCA factors (associated with larger λ_n^{RP}) do not necessarily have highest factor risk premia $\mu_{\Pi_n}^{RP}$. As a result, when $\mu_{\Pi_n}^{RP}$ do not increase adequately with $\sqrt{\lambda_n^{RP}}$, factor prices $SR(\Pi_n^{RP})$ 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 the pricing restriction parameter γ and the combination of RP-PCA and moment fusing help mitigate this issue.

¹⁸Appendix B.1, Equations (B.5) - (B.8), derive the relationships between RP-PCA and moment-fusing covariance matrices in the general setting.

pricing errors across assets), RP-PCA construction prioritizes volatile factors. This is because volatile factors offer a larger share of explained return covariation, once the pricing errors are controlled for in the RP-PCA objective (12). In contrast, in the same limit, ISR-PCA construction prioritizes less volatile (smooth) factors. This is because smooth factors tend to have larger SRs and offer a larger share of explained SRs, once the mean returns are controlled for. This distinction between RP-PCA and ISR-PCA objectives can be quantified by the respective covariance matrices, $\Sigma_x^{RP}(\gamma) \sim [ISRM(X) + (1 + \gamma) \mathbb{1}_{N \times 1} \mathbb{1}_{1 \times N}]$ (see Equation (B.8), Appendix B.1). Indeed, while RP-PCA and ISR-PCA covariance matrices are identical (up to an additional matrix of ones $\mathbb{1}_{N \times 1} \mathbb{1}_{1 \times N}$) in the above limit, RP-PCA ranking (13) and ISR-PCA ranking (11) are opposite to each other (affirming one prioritizes volatile, the other prioritizes smooth, factors).

Relating RP-PCA and moment-fusing constructions: Given a RP-PCA construction governed by the covariance matrix $\Sigma_x^{RP}(\gamma)$ (13), it is instructive to see whether there exists a counterpart moment-fusing construction governed by the same covariance matrix. An application of the moment-fusing design (Proposition 2) provides an affirmative answer to this question. Equation (19) identifies the moment-fusing transformation S , under which the transformed returns Y have the same covariance matrix as $\Sigma_x^{RP}(\gamma)$. Solutions to the sought-after transformed returns Y are,

$$\frac{1}{T} \bar{Y}' \bar{Y} = \underbrace{\frac{1}{T} \tilde{X}' \tilde{X} + (1 + \gamma) \bar{X}' \bar{X}}_{= \Sigma_x^{RP}(\gamma)} \implies Y = XS \quad (21)$$

$$\text{with } S = W_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} \left[\sqrt{\lambda^{RP}} \right] (W^{RP})',$$

where W_x and W^{RP} respectively diagonalize the original and RP-PCA covariance matrices, $W_x' \left(\frac{1}{T} \tilde{X}' \tilde{X} \right) W_x = \text{Diag} [\lambda_x]$, and $(W^{RP})' \Sigma^{RP}(\gamma) W^{RP} = \text{Diag} [\lambda^{RP}]$. Evidently, eigenvalues λ^{RP} carry information about the RP-PCA pricing restrictions (including the first moment $\bar{X}$ of the original returns and pricing restriction parameter γ). After this information is passed on to the transformation S (21), it is fused into the transformed returns $Y = XS$. However, RP-PCA factors $\Pi^{RP} = XW^{RP}$ remain distinct from the moment-fusing factors $\Pi_y = YW^{RP} = XS W^{RP}$.

SR-based prioritization and scaling operations: Recall that scaling (i.e., pure leverage) opera-

tions (3) leave the return space and the underlying pricing model intact. Consider such an operation on a moment-fusing construction, whose initial (before the scaling) factors $\{\Pi_{yn}\}$, $n \in \{1, \dots, N\}$ are traded and pairwise uncorrelated, and span the original asset return space $\{X_n\}$, $n \in \{1, \dots, N\}$. The scaling generates (transformed) factors $\{\Pi_{yn}(\kappa_n) = \kappa_n \Pi_{yn}\}$, $n \in \{1, \dots, N\}$, which remain traded, pairwise orthogonal, and also span $\{X_n\}$. The risk premia, variances, and SRs of the transformed PCs are

$$(22) \quad \mu_{\Pi_{yn}}(\kappa_n) = \kappa_n \mu_{\Pi_{yn}}, \quad \sigma_{\Pi_{yn}}(\kappa_n) = \kappa_n \sqrt{\lambda_{yn}}, \quad SR[\Pi_{yn}](\kappa_n) = SR[\Pi_{yn}].$$

Note that the corresponding original and transformed PCs perfectly correlate, $Corr[\Pi_{yn}, \Pi_{yn}(\kappa_n)] = 1$, have identical prices of risks (SRs), and hence, represent the same risk. Yet, the covariance-based rankings of original $\{\Pi_{yn}\}$ and transformed $\{\Pi_{yn}(\kappa_n)\}$ PCs are not aligned in general due to the presence of free scaling parameters $\{\kappa_n\}$. That is, the covariance-based prioritization in standard PCA is ambiguous under a simple factor scaling, while the SR-based prioritization in the moment-fusing approach is not. Adopting this observation for RP-PCA, we note that scaling operations on RP-PCA factors, $\Pi_n^{RP}(\kappa_n) = \kappa_n \Pi_n^{RP}$, $n \in \{1, \dots, N\}$, do not change the set of these factors.¹⁹ However, the factor scaling changes the RP-PCA factor ranking.²⁰ This observation indicates scaling operations, which are inherent and essential in moment-fusing constructions (Proposition 2), can improve RP-PCA factors when the two approaches are combined.

In summary, RP-PCA and moment-fusing constructions employ respectively pricing restrictions and SR-based objective to identify pricing factors. While the moment-fusing approach is robust to factor scaling transformations, it is not exact because factors explaining a larger shares of SRs of asset returns do not always have larger factor prices. These differences are the basis for improvements in moment-fusing constructions that incorporate RP-PCA's pricing restrictions.

¹⁹To see that a scaling operation leaves the set of RP-PCA solution factors unchanged, note that the transformed factors $\Pi_n^{RP}(\kappa_n) = \kappa_n \Pi_n^{RP}$, $n \in \{1, \dots, N\}$, remain the RP-PCA optimal solution. That is, they diagonalize the respective (transformed) RP-PCA covariance matrix, $\Sigma_{\Pi}^{RP}(\kappa, \gamma) \equiv \frac{1}{T} \tilde{\Pi}^{RP'}(\kappa) \tilde{\Pi}^{RP}(\kappa) + (1 + \gamma) \bar{\Pi}^{RP'}(\kappa) \bar{\Pi}^{RP}(\kappa) = \text{Diag}[\kappa_n^2 \lambda_n^{RP}(\gamma)]$ because $\{\Pi_n^{RP}\}$ diagonalizes the original RP-PCA covariance matrix, $\Sigma_{\Pi}^{RP}(\gamma) \equiv \frac{1}{T} \tilde{\Pi}^{RP'} \tilde{\Pi}^{RP} + (1 + \gamma) \bar{\Pi}^{RP'} \bar{\Pi}^{RP} = \text{Diag}[\lambda_n^{RP}(\gamma)]$.

²⁰This because RP-PCA features a covariance-based prioritization (13) and the RP-PCA factor volatility changes with scaling parameters, $\sigma_{\Pi_{RP}^{xn}}(\kappa_n) = \kappa_n \sqrt{\lambda_{xn}}$. The risk represented by corresponding factors $\Pi_n^{RP}(\kappa_n)$ and Π_n^{RP} remain the same as these factors perfectly correlate.

5 Empirical Analysis: FX Market

This section presents an empirical analysis of the moment-fusing approach and related pricing models in the FX market. Section 5.1 discusses FX data and sources. Section 5.2 presents basic currency strategies and benchmark FX factors from the perspective of a generic currency denomination (numeraire). Section 5.3 presents in- and out-of-sample empirical results for various moment-fusing constructions and benchmark models in different numeraire currencies. Appendix A details the training procedure of pricing models, Appendix B discusses a conceptual relationship between the currency denomination and moment-fusing constructions in the FX market.

5.1 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). When the numeraire (base) currency is USD, the exchange rate notation is such that $S_{i,t}$ USD buys one unit of foreign currency i at time t .

In the original data set from the World Market/Refinitiv (WM/R, or previously, 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 adopted by the Deutsche Bank 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.²¹ We supplement this data set by employing the forward exchange rate discount (hereafter also referred to as the forward discount, constructed as the one-month forward exchange rate minus the spot exchange rate) to estimate the interest rate differential with respect to the U.S. interest rate for the period from May 2020 to October 2025, employing the covered interest rate parity (discussed further in Appendix B.2, Equation (B.10)).

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) share of unexplained (idiosyncratic) return variation. Among these pricing measures, the maximum attainable Sharpe ratio speaks most directly to the moment-fusing constructions for two reasons. Namely, SRs are robust to scaling operations of the type (3), 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 overfit asset returns by delivering high in-sample SRs. Appendix B.3 presents a description of these pricing measures.

5.2 Currency Returns and FX Strategies

This section discusses 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 basic long-short 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. Assuming CIP, this strategy can be implemented using 1-month forward exchange rate contract, generating the following realized return in the USD numeraire (Equation (B.10), Appendix B.2)

$$(23) \quad CT_{i/US}(t + 1) = \ln(S_{i,t+1}) - \ln(F_{i,t}) \equiv s_{i,t+1} - f_{i,t},$$

²¹<http://web.mit.edu/adrienv/www/Data.html>, under the heading “Monthly Changes in Exchange Rates and Global Risk Factors”.

where $s_{i,t+1}$ and $f_{i,t}$ are the natural logarithm respectively of the spot exchange rate in month $t + 1$ and of the 1-month forward exchange rate in month t , between currency i and USD. Since this is a long-short strategy, (23) is an excess return (i.e., return in excess of the US risk-free rate). 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.

Next, we follow Lustig et al. (2011) to construct two prominent factors in the FX market, namely, the Dollar (denoted as RX) and the High-minus-Low (HML) factors. We sort currency returns (23) based on the contemporaneous forward discount $\ln(S_{i,t}/F_{i,t})$ 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,

$$(24) \quad RX(t+1) = \frac{1}{10} \sum_{i=1}^{10} CT_{i/US}(t+1),$$

$$(25) \quad HML(t+1) = \frac{1}{2} \sum_{H_{i,t}=1}^2 CT_{H_{i,t}/US}(t+1) - \frac{1}{2} \sum_{L_{i,t}=1}^2 CT_{L_{i,t}/US}(t+1),$$

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 a strategy borrowing currency B and lending other currencies L from t to $t + 1$ from the perspective of an investor in a base (numeraire) currency H from t to $t + 1$. Generalizing (23), the realized return denominated in currency H of this strategy is,²²

$$(26) \quad CT_{L/B}^H(t+1) = \underbrace{(s_{L/H,t+1} - f_{L/H,t})}_{=CT_{L/H}^H(t+1)} - \underbrace{(s_{B/H,t+1} - f_{B/H,t})}_{=CT_{B/H}^H(t+1)},$$

where $s_{i/H,t} = \ln S_{i/H,t}$ and $f_{i/H,t} = \ln F_{i/H,t}$. Note that the long-short strategy involving cur-

²²Appendix B.2 and the discussion following Equation (B.10) present a derivation of the generic currency return (26).

rencies L , B , and denominated in the numeraire currency H can be decomposed into two basic long-short strategies with offsetting H 's positions, one involving L and H and another involving H and B . When the base and the borrowing currencies coincide with the USD ($B \equiv H \equiv US$), the currency return (26) reduces to (23).

Next, we consider the generic version of the RX strategy above as borrowing (short) the base currency H and lending (long) equally other currencies i 's. The return in the numeraire currency H of this generic RX strategy is

$$(27) \quad RX_H(t+1) = \frac{1}{10} \sum_{i=1}^{10} CT_{i/H}^H(t+1), \quad \text{where} \quad CT_{i/H}^H(t+1) = s_{i/H,t+1} - f_{i/H,t}.$$

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

$$(28) \quad HML_{H,t+1} = \frac{1}{2} \sum_{H_{i,t}=1}^2 CT_{H_{i,t}/H}^H(t+1) - \frac{1}{2} \sum_{L_{i,t}=1}^2 CT_{L_{i,t}/H}^H(t+1),$$

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 (25) and in the base of currency H

(28) are essentially the same.²³ 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 approach is a generic approach to construct pricing factors. For the FX market, we extend the moment-fusion constructions of SR-PCA (7), ISR-PCA (10) by optimally incorporating the pricing restrictions as described in the extended constructions (14) and (15), Section 3.3. Presenting an empirical analysis of these extended moment-fusion constructions allows us to assess the contributions of moment-fusing and pricing restrictions to an optimal overall pricing performance. In particular, the time series plot of optimal pricing restriction parameter γ (wherein $\gamma = -1$ signifies the absence of the pricing restrictions) indicates the influence of the pricing restrictions dynamically. Our analysis is performed separately for 4 different representative numeraires (USD, JPY, AUD, GBP) of the developed currencies in the sample. 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 $\{CT_{i/H}^H\}$ (27), where i is in the set of $N = 10$ currencies (excluding the current base currency H) in the sample. We then construct the transformed currency returns by incorporating the first moment m_x of the original currency returns $\{CT_{i/H}^H\}$, $i \in \{1, \dots, N\}$, using the extended moment-fusing constructions SR-PCA or ISR-PCA (14). 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. To compare with the two FX benchmark factors RX (27) and HML (28), 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 (15). 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 restriction parameter γ dynamically (reported in Figure 1 for the representative USD numeraire).

²³In 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.

5.3 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).²⁴ 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 (12) 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

²⁴That is, the top two (1st and 2nd) factors are ranked based on model-specific priority criterion of the model under consideration.

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 (27), 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 returns are measured in different base currencies, so their correlations include an inherent exchange rate component between currency bases. Instead, the cross-numeraire correlations of HML present a measure for this exchange rate component because HML is largely the same strategy (borrowing low, lending high interest rate currencies) across bases as observed below (28). 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 B.3), of 7 different factor models in the FX market. The models are Lustig et al. (2011)'s RX & HML (27)-(28), PCA, Lettau and Pelger (2020)'s RP-PCA (13), extended moment-fusing models SR-PCA and ISR-PCA (14), and combined moment-fusing models Π_{ew} (equally weighted) and Π_{ow} (optimally weighted) (15). 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

parameter γ (Appendix A.1).

In sample, the benchmark RX & HML model (together with the combined moment-fusing model Π_{ow} and the extended moment-fusing model ISR-PCA) delivers the highest maximum attainable (annualized) Sharpe ratio of 0.39 in the USD numeraire, confirming Lustig et al. (2011)'s findings on the superior pricing performance of the benchmark model in USD. For the numeraire of JPY, AUD, and GBP respectively, RX & HML model delivers an annualized Sharpe ratio of 0.43, 0.35, and 0.33 and remains among top-performing models in Table 3. For these 3 numeraire currencies, the model Π_{ow} respectively delivers a higher annualized Sharpe ratio of 0.55, 0.40, and 0.37 and is the top-performing model consistently for all considered numeraire currencies. The moment-fusing model ISR-PCA delivers significant maximum attainable Sharpe ratios, but also large shares of unexplained return variation σ_e . It reflects the reverse eigenvalue ranking (i.e., focusing less on the pricing measure of σ_e) in this model. RP-PCA performs similarly to PCA because as noted above, pricing restrictions have a subdued (in-sample) 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 and pricing restrictions have important impacts. Extended moment-fusing factors and related models consistently deliver highest OOS maximum attainable Sharpe ratios. First, ISR-PCA and Π_{ow} continue to outperform OOS, delivering a Sharpe ratio of 0.28, 0.72, 0.33 and 0.65, and 0.28, 0.70, 0.81, and 0.65, respectively in the numeraire of USD, JPY, AUD and GBP. Model SR-PCA delivers sub-dominant but significant OOS Sharpe ratio of 0.09, 0.36, 0.25, and 0.23 respectively in these four numeraires, while the benchmark model RX & HML delivers 0.09, 0.18, 0.01, and 0.11 respectively. PCA and RP-PCA differ significantly OOS only in the GBP numeraire. Finally, Figure 1 reports the incorporation and impact of the pricing restrictions on moment-fusing constructions in the USD numeraire. Such an incorporation is characterized by the monthly time series of parameter γ in the objective function (12), wherein $\gamma = -1$ signifies the absence of pricing constraints. For all three models RP-PCA, SR-PCA, ISR-PCA, γ has the value of -1 for most of months in the time series, indicating that moment-fusing FX factors (from the perspective of the USD denomination) can be identified and constructed without invoking the pricing restrictions.

In summary, the benchmark RX & HML model's pricing performance varies significantly with the numeraire currency. Overall, the moment-fusing models ISR-PCA and Π_{ow} are the outperforming model both in- and out-of-sample and for most numeraire currencies. These results show the importance of moment-fusing and pricing restrictions in identifying robust FX pricing factors.

6 Empirical Analysis: Equity Market

This section presents an empirical analysis of the moment-fusing approach and related pricing models in the equity market. Section 6.1 discusses equity data and sources. Section 6.2 presents in- and out-of-sample empirical results for various moment-fusing constructions and benchmark models in various equity data sets.

6.1 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.²⁵ While the sourced data (Footnote 25) 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).²⁶ 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 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.

²⁵ The characteristics are available at <https://www.serhiykozak.com/data>, under the heading "Portfolio Sorts", and discussed in Kozak et al. (2018).

²⁶ 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>

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 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 restriction parameter γ dynamically. Presenting an empirical analysis of these extended moment-fusing constructions allows us to assess the contributions of moment fusing and pricing restrictions to an optimal overall pricing performance. The time series plot of optimal parameter γ (reported in Figure 2) indicates the influence of the pricing restrictions dynamically.

6.2 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 B.3). The nomenclature is in Table 1.

[Table 4]

Table 4 reports the pricing performance, quantified by 4 pricing measures (Appendix B.3), of 6 different factor models using 74 extreme-decile anomaly portfolios. The models are PCA, Lettau and Pelger (2020)'s RP-PCA (13), extended moment-fusing models SR-PCA and ISR-PCA (14), and combined moment-fusing models Π_{ew} (equally weighted) and Π_{ow} (optimally weighted)

(15). 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 parameter γ (Appendix A.1).

In sample, for $K = 3$ factors, the moment-fusing models Π_{ow} and SR-PCA, and the benchmark model RP-PCA are top-performing models, delivering a maximum attainable (monthly) Sharpe ratio of 0.45, 0.45, and 0.37 respectively. For $K = 5$ factors, the same models remain top three models, delivering a (monthly) Sharpe ratio of 0.49, 0.49, and 0.45 respectively. Note that these Sharpe ratios (after being converted to annualized values) are significantly higher than those obtained using FX data (Section 5.3) in part because a larger number of returns in the equity data sample offers a higher degree of diversification. Out of sample, the three models deliver a maximum attainable (monthly) Sharpe ratio of 0.09, 0.09, and 0.10 for $K = 3$ factors, and 0.55, 0.55, and 0.45 for $K = 5$ factors, and remain top-performing models. The OOS out-performance of Π_{ow} compared to Π_{ew} (for both $K = 3$ and $K = 5$) shows the robustness of optimally choosing the weight between components in the combined moment-fusing model (15). The moment-fusing model ISR-PCA delivers a significant maximum attainable Sharpe ratio OOS for both $K = 3$ and $K = 5$ retained factors, but also have larger share 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 $K = 5$ retained factors, the moment-fusing models Π_{ow} and SR-PCA continue to dominate in terms of the maximum attainable Sharpe ratio, delivering a (monthly) value of 0.40 (for $K = 3$) and 0.44 (for $K = 5$), while the benchmark model RP-PCA delivers respective 0.24 and 0.34. Out of sample, for $K = 3$ retained factors, RP-PCA delivers a maximum attainable (monthly) Sharpe ratio of 0.19, dominating the corresponding value of 0.16 for the models Π_{ow} and SR-PCA. For $K = 5$ retained factors, the moment-fusing Π_{ow} and SR-PCA and the benchmark RP-PCA are the top-performing models, delivering the highest OOS

maximum attainable (monthly) Sharpe ratio of 0.37. These patterns again show the consistency between in- and out-of-sample results, indicating the robustness of factors constructed using moment fusing and pricing restrictions.

[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 Π_{ow} is the out-performing model, delivering a maximum attainable (monthly) Sharpe ratio of 0.34 for $K = 3$ and 0.37 $K = 5$ retained factors. The moment-fusing SR-PCA and the benchmark RP-PCA are next, delivering respectively a (monthly) Sharpe ratio of 0.25 and 0.26 for $K = 3$, and 0.37 and 0.36 for $K = 5$. Out of sample, the moment-fusing model ISR-PCA dominates for both $K = 3$ and $K = 5$ retained factors, delivering a maximum attainable (monthly) Sharpe ratio of 0.42 (for $K = 3$) and 0.51 (for $K = 5$). The benchmark model RP-PCA delivers respectively 0.25 and 0.34, followed by the moment-fusing models Π_{ow} and SR-PCA delivering 0.24 and 0.27. Finally, Figure 2 reports the incorporation and impact of the pricing restrictions on moment-fusing constructions for the 74 extreme-decile anomaly portfolios. Such an incorporation is characterized by the monthly time series of parameter γ in the objective function (12), wherein $\gamma = -1$ signifies the absence of pricing constrains. For all three models RP-PCA, SR-PCA, ISR-PCA, $\gamma \neq -1$ for many months in the time series, indicating that pricing restrictions are important in the identification and construction of moment-fusing equity factors (in the LP74 data set), in contrast with significantly smaller pricing restriction influence on the FX factor construction (Figure 1).

In summary, the combined moment-fusing factor construction (with optimal weight) Π_{ow} is the overall outperforming pricing mode consistently for most of in- and out-of-sample configurations of equity data sets and retained factors. The benchmark RP-PCA and the moment-fusing SR-PCA perform similarly in most of configurations. For the smaller equity data set FF25, the moment-fusing model ISR-PCA is the out-performing model OOS. However, ISR-PCA also has a slightly elevated share of unexplained return variation σ_e for FF25, which is consistent with the reverse eigenvalue ranking in this model. The overall findings show important improvements in equity pricing factors whose constructions are fused with original return moments and extended

with pricing restrictions.

7 Conclusions

The current paper introduces the moment-fusing approach to identify important common risk factors that price asset returns in the cross section and out of sample. The approach 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 that are informed by the risk premia and Sharpe ratios of original returns while also utilizing the elegance and power of the PCA methodology. The moment-fusing approach is extendable and can be improved upon by incorporating prominent features of literature's existing pricing frameworks.

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 in and out of sample. The empirical analysis shows that pricing restrictions complement the moment-fusing approach in furthering the performance of pricing factors, though their effect is more significant in equity market than in FX market.

The moment-fusing constructions are both general and flexible. While the current paper focuses on fusing information about SRs and optimizing factor prices, the moment-fusing approach can also fuse other pricing (e.g., skewness, kurtosis, risk loadings and pricing errors) and economic (macro and stock characteristics) information and optimize other desired moments of the resulting factors.

References

- Bai, J., 2003, Inferential Theory for Factor Models of Large Dimensions, *Econometrica* 71, 135–171.
- Barillas, F., and J. Shanken, 2017, Which Alpha?, *Review of Financial Studies* 30, 1316–1338.
- Bessembinder, H., A. Burt, and C. Hrdlicka, 2022, Time Series Variation in the Factor Zoo https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3992041.
- Bessembinder, H., A. Burt, and C. Hrdlicka, 2025, Can We Tame the Factor Zoo? The Roles of Alternative Databases and Extraction Methods https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5150471.
- 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, 5449–5956.
- Maurer, T., T. Tô, and N-K. Tran, 2022, Pricing Implications of Covariances and Spreads in Currency Markets, *The Review of Asset Pricing Studies* 12, 336–388.
- Maurer, T., T. Tô, and N-K. Tran, 2023, Market Timing and Predictability in FX Markets, *Review of Finance* 27, 223–246.
- 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 models, data sets, and pricing measures employed in the paper.

Panel A	Factor Models
PCA	Principal component analysis without pricing restrictions.
RX & HML	RX is the dollar factor and HML is the carry trade factor Lustig et al. (2011) .
RP-PCA	Risk-Premium Principal Component Analysis: PCA with pricing restrictions with pricing restriction parameter $(1 + \gamma)$ Lettau and Pelger (2020) .
SR-PCA	Sharpe Ratio PCA: Returns are transformed so that their variances equal SRs of original returns, extended with pricing restrictions, before running PCA.
ISR-PCA	Inverse Sharpe Ratio PCA: Returns are transformed to have an identical mean, extended with pricing restrictions, before running PCA.
Π_{ew}	Equally weighted portfolio combining corresponding factors of ISR-PCA and SR-PCA.
Π_{ow}	Optimally weighted portfolio combining corresponding factors of ISR-PCA and SR-PCA.
Panel B	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 C	Pricing Measures
SR	The maximum Sharpe ratio attained with the factors
$RMSE_{\alpha}$	The square root of the average squared pricing error
$\bar{\sigma}_e$	The share (%) of unexplained return variation
GRS	Gibbons-Ross-Shanken 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	$RMSE_{\alpha}$	$\bar{\sigma}_e$	GRS	p		SR	$RMSE_{\alpha}$	$\bar{\sigma}_e$	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	0.39	0.01	34.55	0.50	0.89	ISR-PCA	0.28	0.01	38.87	0.56	0.85
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}	0.24	0.01	25.98	0.35	0.97	Π_{EW}	0.11	0.01	25.25	0.48	0.90
Π_{OW}	0.39	0.01	34.55	0.50	0.89	Π_{OW}	0.28	0.01	38.87	0.56	0.85
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	0.53	0.01	23.89	0.24	0.99	ISR-PCA	0.72	0.01	23.42	0.61	0.81
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}	0.25	0.01	17.74	0.41	0.94	Π_{EW}	0.44	0.01	17.57	0.34	0.97
Π_{OW}	0.55	0.01	24.50	0.19	1.00	Π_{OW}	0.70	0.01	25.07	0.70	0.72
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	0.13	0.01	25.98	0.56	0.85	ISR-PCA	0.33	0.01	32.33	0.53	0.87
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}	0.39	0.01	23.84	0.16	1.00	Π_{EW}	1.13	0.02	31.12	1.25	0.26
Π_{OW}	0.40	0.00	24.17	0.10	1.00	Π_{OW}	0.81	0.01	31.54	0.81	0.62
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	0.19	0.01	42.16	0.90	0.54	ISR-PCA	0.65	0.01	44.35	0.49	0.90
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}	0.37	0.01	37.17	0.18	1.00	Π_{EW}	0.65	0.01	38.04	0.40	0.94
Π_{OW}	0.37	0.01	37.17	0.18	1.00	Π_{OW}	0.65	0.01	38.04	0.40	0.94

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 restriction parameter 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	$RMSE_{\alpha}$	$\bar{\sigma}_e$	GRS	p		SR	$RMSE_{\alpha}$	$\bar{\sigma}_e$	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	0.25	0.31	19.97	5.87	0	ISR-PCA	0.18	0.34	20.19	5.31	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}	0.42	0.26	17.87	4.51	0	Π_{EW}	0.02	0.36	18.95	5.78	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	0.33	0.29	16.43	5.32	0	ISR-PCA	0.30	0.25	15.36	4.60	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	0.22	14.63	3.95	0	Π_{EW}	0.22	0.24	14.31	5.40	0
Π_{OW}	0.49	0.24	12.86	3.80	0	Π_{OW}	0.55	0.23	13.71	5.10	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 restriction parameter 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	$RMSE_{\alpha}$	$\bar{\sigma}_e$	GRS	p		SR	$RMSE_{\alpha}$	$\bar{\sigma}_e$	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	0.27	0.16	16.17	2.00	0	ISR-PCA	0.21	0.17	16.07	3.02	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}	0.36	0.15	15.61	1.85	0	Π_{EW}	0.10	0.17	16.41	3.07	0
Π_{OW}	0.40	0.18	15.19	1.81	0	Π_{OW}	0.16	0.18	16.28	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	0.28	0.15	14.61	1.99	0	ISR-PCA	0.26	0.15	13.34	2.90	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}	0.38	0.15	13.17	1.81	0	Π_{EW}	0.10	0.15	13.21	3.18	0
Π_{OW}	0.44	0.17	12.28	1.74	0	Π_{OW}	0.37	0.15	12.95	3.03	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 restriction parameter 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	$RMSE_{\alpha}$	$\bar{\sigma}_e$	GRS	p		SR	$RMSE_{\alpha}$	$\bar{\sigma}_e$	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	0.26	0.15	6.68	3.98	0	ISR-PCA	0.42	0.13	9.22	4.40	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}	0.31	0.22	12.09	3.18	0	Π_{EW}	0.25	0.16	7.79	5.76	0
Π_{OW}	0.34	0.12	10.85	2.69	0	Π_{OW}	0.25	0.16	8.01	5.57	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	0.29	0.13	4.74	3.51	0	ISR-PCA	0.51	0.12	6.01	3.91	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}	0.37	0.12	5.75	2.11	0	Π_{EW}	0.28	0.12	5.08	3.88	0
Π_{OW}	0.37	0.10	5.47	1.98	0	Π_{OW}	0.27	0.11	5.18	3.48	0

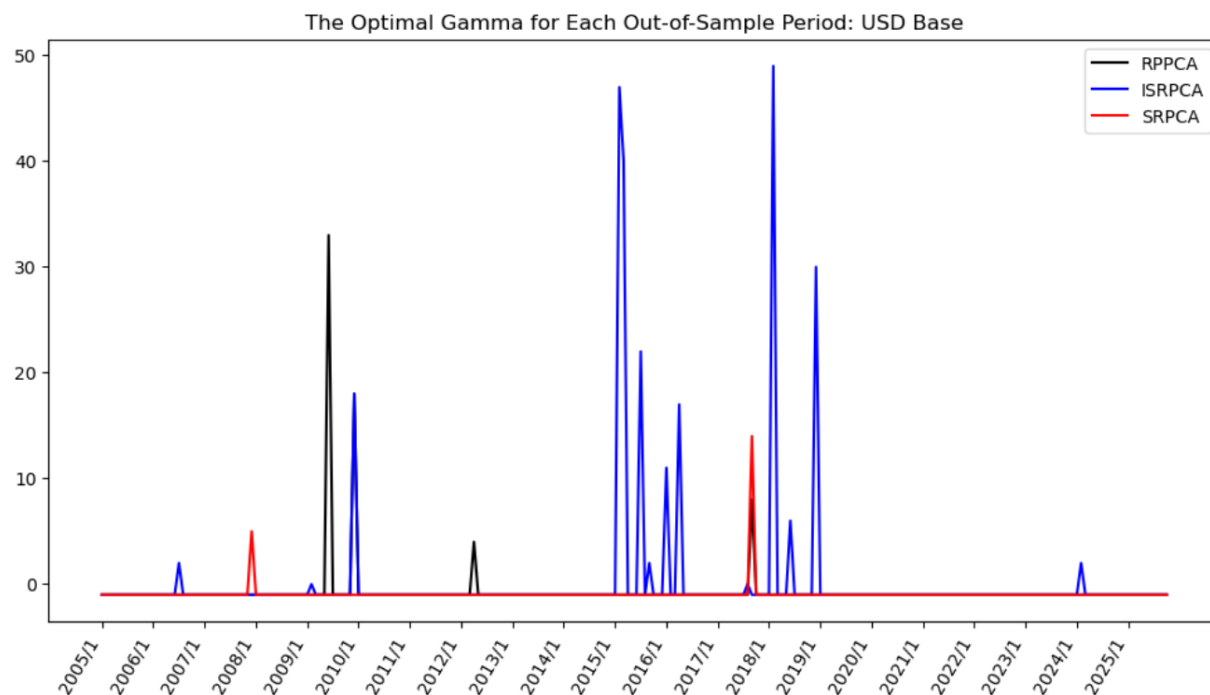


Figure 1: **The Time Series of the Optimal γ of the Currency Carry Trade Return Data: USD Base.** This plot shows the time series of the optimal pricing restriction parameter γ (12) and (14) in each out-of-sample period that leads to the optimal Sharpe ratios for the USD base.

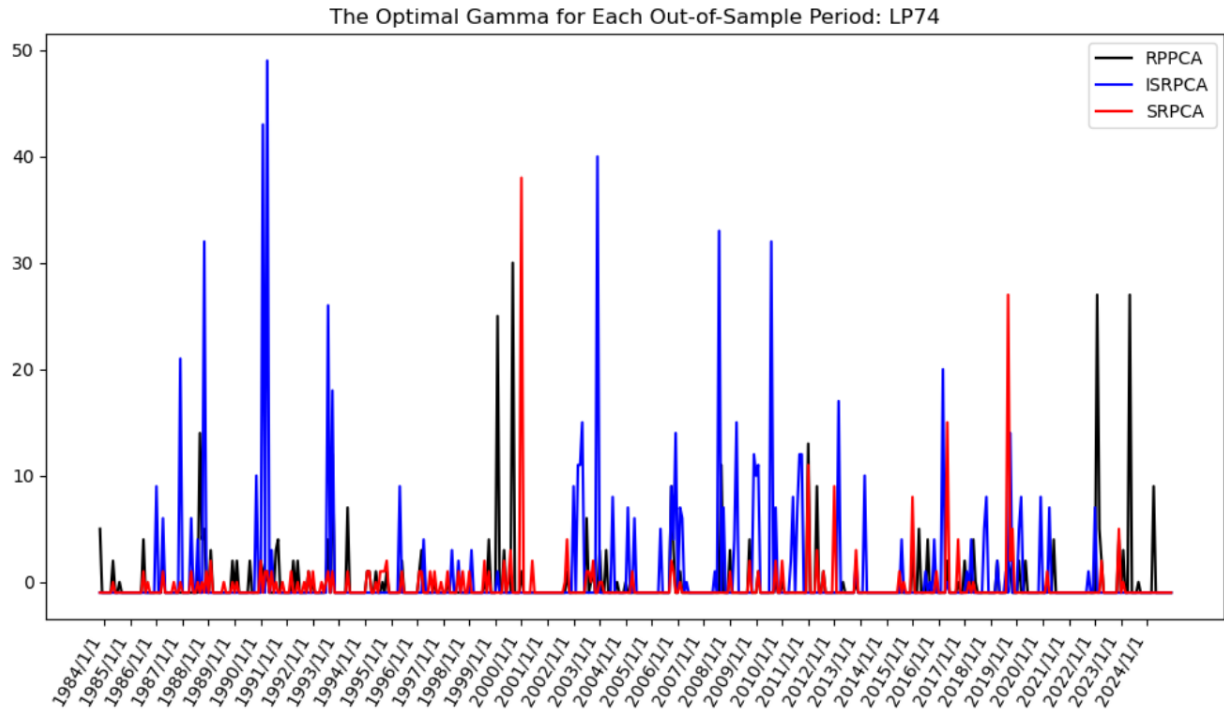


Figure 2: **The Time Series of the Optimal γ of the Stock Portfolio Return Data: LP74.** This plot shows the time series of the optimal pricing restriction parameter γ (12) and (14) in each out-of-sample period that leads to the optimal Sharpe ratios for the LP74 test assets.

Online Appendices

to the manuscript **Moment Fusing: An Informed Construction of Pricing Factors.**

These online appendices present supporting details omitted in the main text: Appendix [A](#) concerns empirical aspects and Appendix [B](#) concerns quantitative aspects.

A Further Empirical Details

This appendix provides further details for the empirical implementation and analysis in the main text: Appendix [A.1](#) summarizes the training steps for pricing models, Appendix [A.2](#) describes the combination of different pricing models. Appendix [A.3](#) lists characteristics concerning equity portfolios.

A.1 The Training Procedure for the Pricing Restriction Parameter

We consider a factor model listed in Table [1](#), referring to it generically as factor model F with K factors ($K = 5$ for the equity return data and $K = 2$ for the currency carry trade return data). Recall that for the extended moment-fusing constructions, we incorporate [Lettau and Pelger \(2020\)](#)'s pricing restrictions into the moment-fusing approach, employing a parameter γ to regularize the cross-sectional pricing errors and integrate it with the standard PCA as specified in Equation [\(12\)](#). We employ a training procedure to determine the pricing restriction parameter γ dynamically in expansive windows. The training procedure is as follows.

1. We begin with an initial 240-month training window ($t = -239$ to $t = 0$) and the associated asset returns X . We use the information from the initial 240-month training window and return data to identify an optimal pricing restriction parameter γ which maximizes the in-sample attainable Sharpe ratio of the factor model F .^{[1](#)}

¹When pricing restrictions are incorporated, optimal factors as well as the maximum attainable Sharpe ratio delivered by the resulting factor model are function of the pricing restriction parameter γ , allowing γ to be a control variable in the maximization of the model's Sharpe ratio. The pricing restriction parameter is searched over the grid $[-1, 15]$ for LP74, $[-1, 10]$ for LP370, and $[-1, 50]$ for FF25 and all four FX data sets. The difference in search grids is due to different computational constraints in different data sets. Moreover, for all data sets and all models, results are robust to different upper limits for γ as long as these limits exceed 10.

2. We form the out-of-sample factor returns in month $t = 1$ by employing out-of-sample asset returns in month $t = 1$ and the factor loadings $W(t = -239 \rightarrow t = 0)$ in the training window.² The OOS factor returns in month $t = 1$ are, $F_{OOS,t=1} = X(t = 1) W(t = -239 \rightarrow t = 0)$. These factor returns then expand the training window by one month ($t = -239$ to $t = 1$) and determine the optimal factor return of the model F .³ We record this optimal factor return as of month $t = 1$.
3. Next, we repeat the above steps: (i) identifying the updated optimal pricing restriction parameter for the 241-month training window (by grid search), (ii) forming out-of-sample factor returns in month $t = 2$ by employing the OOS asset returns $X(t = 2)$ in month $t = 2$ and the in-sample (i.e., in the training window) factor loadings $W(t = -239 \rightarrow t = 1)$, and (iii) determining and recording the optimal factor return of the model F as of month $t = 2$.
4. We repeat this process until the end of the observations in the sample. As a result of this procedure, we obtain the following time series of the time length $T - 240$, where T is the total number of observations in the expansive windows.
 - The time series of the (updated) pricing restriction parameter γ .
 - The time series of the (recorded) optimal factor return of the model F .

We determine the Sharpe ratio using the recorded optimal factor return time series and employ it to proxy for the OOS maximum attainable Sharpe ratio estimate of the factor model F .

A.2 Combining Factor Models

This appendix describes the procedure to combine the SR-PCA and ISR-PCA models. As we extend these moment-fusing models by pricing restrictions, the asset return inputs are transformed asset returns Y (which are transformed from original asset returns X using respectively

² Recall that when factors are constructed as principal components in the RP-PCA approach, they are linear in asset returns $F = XW$, where the loading matrix W diagonalizes the return covariance matrix adjusted by pricing restrictions, $W' \left[\frac{1}{T} \tilde{X}' \tilde{X} + (1 + \gamma) \bar{X}' \bar{X} \right] W = \text{Diag}(\Lambda)$. That is, given return data X and a parameter γ , loadings W are solved and obtained as eigenvectors of the risk-premium adjusted covariance matrix $\frac{1}{T} \tilde{X}' \tilde{X} + (1 + \gamma) \bar{X}' \bar{X}$. As a result, the loading matrix is a function of the pricing restriction parameter, $W = W(\gamma)$. In the case of OOS transformed asset returns, we replace X by the OOS transformed returns $Y = XS(m)$, where the X represent the OOS original returns, $S(m)$ the in-sample moment-fusing transformation matrix (determined in the training window).

³ Optimal factor return is the return on the tangency portfolio, whose weights $\Sigma_F^{-1} \mu'_F$ are computed using data in the newly expanded training window from $t = -239$ to $t = 1$.

SR-PCA (6) and ISR-PCA (9) transformations, see also Footnote 2). We apply the weight of w_1 to ISR-PCA model and $1 - w_1$ to SR-PCA model to combine them into a new model, whose factors are

- Π_{ew} : applying (equal) weights $\{w_1 = 0.5, 1 - w_1 = 0.5\}$ respectively to ISR-PCA and SR-PCA factors.
- Π_{ow} : applying (optimal) weights $\{w_1, 1 - w_1\}$ respectively to ISR-PCA and SR-PCA factors, where the determination of the (optimal) combination weight w_1^* is carried jointly with the determination of the combined model's in-sample and out-of-sample Sharpe ratios (detailed below).

The **in-sample Sharpe ratio** of the combined portfolio from the ISR-PCA and SR-PCA factors, which involves the full sample (and no training for the pricing restriction parameter), is determined as follows. First, we implement two separate searches, one for the optimal pricing restriction $\gamma_{ISR-PCA}$ associated with ISR-PCA model, the other for γ_{SR-PCA} associated with SR-PCA, in the parameter grid of $[-1, 50]$. The optimization objective is to maximize in-sample Sharpe ratios (respectively for ISR-PCA and SR-PCA models) for the full sample period. The equally weighted combined factor model is constructed by applying $w_1 = 0.5$ to ISR-PCA and $1 - w_1 = 0.5$ to SR-PCA, or $\Pi_{ew} = 0.5F_{ISR-PCA} + 0.5F_{SR-PCA}$. To identify the weight w_1^* in the optimally weighted combined model, $\Pi_{ow}(w_1) = w_1F_{ISR-PCA} + (1 - w_1)F_{SR-PCA}$, we employ a grid search of the range $[-2, 2]$ for w_1 with the objective to maximize the in-sample Sharpe ratio of the combined model. The solution Sharpe ratio (associated with the optimal solution w_1^*) proxies for the maximum attainable in-sample Sharpe ratio estimate of the combined model.

The **out-of-sample Sharpe ratio** of the combined portfolio from the ISR-PCA and SR-PCA factors is determined in conjunction with the a training procedure for parameters in expansive windows. This procedure is described in Appendix A.1 and employed here with minor modifications. To simplify the search while reducing computational cost, we employ the same pricing restriction parameter γ for both ISR-PCA and SR-PCA models and combine their respective factors.⁴ First, in the initial 240-month training window, we parameterize factors of each component (ISR-PCA or SR-PCA) model by a common pricing restriction parameter γ , before com-

⁴This can be seen as an additional constraint for a uniform pricing restriction of the combined model (while

binning these factors using weights w_1 and $1 - w_1$ respectively for ISR-PCA and SR-PCA. We then search for the numerical values of these parameters (i.e., grid searches) that maximize the attainable Sharpe ratio of the combined model in the training window (see Footnotes 1 and 2). Next, we construct the combined factor returns in month $t = 1$ by employing OOS asset returns in month $t = 1$ and the factor loadings in the training window. These combined factor returns then expand the training window by one month ($t = -239$ to $t = 1$) and determine the optimal factor return of the combined model. We record this OOS optimal factor return of the combined model. Repeating this process and updating the grid searches in expansive windows generates a time series of the OOS optimal factor return of the combined model. As before, we compute the Sharpe ratio using this return time series to proxy for the OOS maximum attainable Sharpe ratio estimate of the combined model.

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 re-

the flexibility in combining ISR-PCA and SR-PCA components remains via the presence of an optimal combination weight w_1). As a result of this constraint, the obtained maximum attainable Sharpe ratio of the combined model is a conservative estimate.

lated 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).

B Further Quantitative Details

This appendix provides further quantitative details for the analysis in the main text: Appendix [B.1](#) presents technical derivations and proofs omitted in the main text. Appendix [B.2](#) discusses the relationship between the currency denomination's choice and moment-fusing constructions. Appendix [B.3](#) describes four pricing measures employed in the paper.

B.1 Derivations and Proofs

Proof of Proposition 1 (Sharpe Ratio rotational invariance): Consider any two equivalent (original and transformed) risky return bases $(\{X_n\}_{n=1}^N$ and $\{Y_n\}_{n=1}^N$) that are related by a $N \times N$ transformation matrix S of full rank $Y_{T \times N} = X_{T \times N} S_{N \times N}$. We implement PCA on the original and transformed return bases by diagonalizing respective covariance matrices, obtaining

respective sets of original and transformed PCs as columns $\Pi_x = XW_x$ and $\Pi_y = YW_y$ (as discussed above (B.2)).

To relate original and transformed PCs and their SRs, we first employ the pure leverage operation (3) to standardize these PCs, $\hat{\Pi}_x \equiv \Pi_x \text{Diag}\left(\frac{1}{\sqrt{\lambda_x}}\right)$ (B.3) and $\hat{\Pi}_y \equiv \Pi_y \text{Diag}\left(\frac{1}{\sqrt{\lambda_y}}\right)$ (B.2). They constitute two orthonormal bases of the asset return space, hence can be rotated from one to the other by a $N \times N$ rotation matrix $\hat{\Pi}_y = \hat{\Pi}_x Q$, or equivalently, $\hat{\Pi}_y' = Q' \hat{\Pi}_x'$ as explained above (B.4). Next, taking the mean of every row in the last matrix equation, $\bar{\hat{\Pi}}_y' = Q' \bar{\hat{\Pi}}_x'$, and noting that the means of standardized PCs are SRs of PCs, we have,⁵

$$(B.1) \quad \underbrace{\begin{bmatrix} \frac{\bar{\Pi}_{y1}}{\sqrt{\lambda_{y1}}} \\ \vdots \\ \frac{\bar{\Pi}_{yN}}{\sqrt{\lambda_{yN}}} \end{bmatrix}}_{=SR[\Pi_y]} = Q' \underbrace{\begin{bmatrix} \frac{\bar{\Pi}_{x1}}{\sqrt{\lambda_{x1}}} \\ \vdots \\ \frac{\bar{\Pi}_{xN}}{\sqrt{\lambda_{xN}}} \end{bmatrix}}_{=SR[\Pi_x]}, \quad \text{or,} \quad \begin{cases} SR[\Pi_y] = Q' SR[\Pi_x], \\ SR[\Pi_x] = Q SR[\Pi_y]. \end{cases}$$

These relationships establish the Sharpe ratio rotational invariance underlying Proposition 1 ■

Proof of Proposition 2 (Moment-fusing Design): The proof has four instructive steps. In the first step, given the transformed covariance matrix Σ_y , we diagonalize it to obtain the associated eigenvalues $\{\lambda_{yn}\}_{n=1}^N$ and identify them with the variances of the transformed PC factors, $\frac{1}{T} \tilde{\Pi}'_{yn} \tilde{\Pi}_{yn} = \lambda_{yn}$, $n \in \{1, \dots, N\}$, or $\frac{1}{T} \tilde{\Pi}'_y \tilde{\Pi}_y = \text{Diag}[\lambda_y]$ in matrix form.⁶ We can perform pure leverage operations to standardize the transformed factors, obtaining $\{\hat{\Pi}_y\}_{n=1}^N$ of unit variances,

$$(B.2) \quad \hat{\Pi}_y = \Pi_y \text{Diag}\left[\frac{1}{\sqrt{\lambda_y}}\right] \implies \frac{1}{T} \tilde{\Pi}'_y \tilde{\Pi}_y = \mathbb{I}_{N \times N}.$$

In the second step, given the original asset returns, we implement the PCA on the original assets, obtaining original factors $\Pi_x = XW_x$ and their variances $\frac{1}{T} \tilde{\Pi}'_x \tilde{\Pi}_x = \text{Diag}[\lambda_x]$ and their standard-

⁵To compute the mean of every row of the matrix $\hat{\Pi}_y' = Q' \hat{\Pi}_x'$, we multiply to the left of both sides by $\frac{1}{T} \mathbb{1}_{T \times 1}$, obtaining $\bar{\hat{\Pi}}_y' = Q' \bar{\hat{\Pi}}_x'$, where the $N \times 1$ column $\bar{\hat{\Pi}}_x'$ contains means of N standardized PCs $\{\hat{\Pi}_{xn}\}_{n=1}^N$, and $\bar{\hat{\Pi}}_y'$ contains means of N standardized PCs $\{\hat{\Pi}_{yn}\}_{n=1}^N$. Since these standardized PCs are constructed as $\hat{\Pi}_x = \Pi_x \text{Diag}\left(\frac{1}{\sqrt{\lambda_x}}\right)$, their means are $\bar{\hat{\Pi}}_{xn} = \frac{\bar{\Pi}_{xn}}{\sqrt{\lambda_{xn}}}$ which are the SRs of PCs Π_{xn} , $n \in \{1, \dots, N\}$. Similarly, the means of standardized PCs $\hat{\Pi}_{yn}$ are the SRs of PCs Π_{yn} , $n \in \{1, \dots, N\}$. These results deliver (B.1).

⁶Recall from a standard PCA process that the variances of PCs are eigenvalues of the covariance matrix, $\frac{1}{T} \tilde{\Pi}'_y \tilde{\Pi}_y = W_y' \Sigma_y W_y = \text{Diag}[\lambda_y]$, or $\text{Var}(\Pi_{yn}) = \lambda_{yn}$.

ized version $\{\hat{\Pi}_x\}_{n=1}^N$ of unit variances,

$$(B.3) \quad \hat{\Pi}_x = \Pi_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] \implies \frac{1}{T} \tilde{\Pi}'_x \tilde{\Pi}_x = \mathbb{1}_{N \times N}.$$

In the third step, note that original and transformed PCs are non-redundant and traded excess returns. Therefore, the standardized factors $\{\hat{\Pi}_{yn}\}_{n=1}^N$ (B.2) and $\{\hat{\Pi}_{xn}\}_{n=1}^N$ (B.3) are two equivalent orthonormal bases of the same return space. Since any two equivalent orthonormal bases are related by a rotation, we have $\hat{\Pi}_y = \hat{\Pi}_x Q$ where Q is a $N \times N$ orthogonal (rotation) matrix, $Q'Q = QQ' = \mathbb{1}_{N \times N}$. The transformed asset returns $\{Y_n\}_{n=1}^N$ whose covariance matrix is the given Σ_y then can be identified by tracing and reversing the sequence of all transformations involved,⁷

$$(B.4) \quad Y \equiv XW_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} [\sqrt{\lambda_y}] W'_y.$$

In the final (verification) step, we compute the covariance matrix of these transformed returns,

$$\begin{aligned} \frac{1}{T} \tilde{Y}' \tilde{Y} &= \frac{1}{T} \left(\tilde{X} W_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} [\sqrt{\lambda_y}] W'_y \right)' \left(\tilde{X} W_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} [\sqrt{\lambda_y}] W'_y \right) \\ &= W_y \text{Diag} [\lambda_y] W'_y = \Sigma_y. \end{aligned}$$

This verification step shows that Q can be any $N \times N$ orthogonal matrix, which implies that there are multiple transformed return solutions Y to the given covariance structure $\frac{1}{T} \tilde{Y}' \tilde{Y} = \Sigma_y$. Note that Equation (B.4) delivers the transformation (19) of Proposition 2 ■

Relationships between RP-PCA and SR-PCA covariance matrices: Using relation (7), we can express the original covariance matrix Σ_x in terms of the SR-PCA covariance matrix $SRM(X)$,

$$\Sigma_x = \text{Diag} \left[\frac{|\sigma_{xn}|^{\frac{3}{2}}}{(\mu_{xn})^{\frac{1}{2}}} \right] SRM(X) \text{Diag} \left[\frac{|\sigma_{xn}|^{\frac{3}{2}}}{(\mu_{xn})^{\frac{1}{2}}} \right].$$

Substituting the above expression of Σ_x into the RP-PCA covariance matrix (13) yields the the

⁷Specifically, the sequence is, $Y = \Pi_y W'_y = \hat{\Pi}_y \text{Diag} [\sqrt{\lambda_y}] W'_y = \hat{\Pi}_x Q \text{Diag} [\sqrt{\lambda_y}] W'_y = \Pi_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} [\sqrt{\lambda_y}] W'_y = XW_x \text{Diag} \left[\frac{1}{\sqrt{\lambda_x}} \right] Q \text{Diag} [\sqrt{\lambda_y}] W'_y$.

relationship between RP-PCA and SR-PCA covariance matrices,

$$(B.5) \quad \Sigma_x^{RP}(\gamma) = \text{Diag} \left[\frac{|\sigma_{xn}|^{\frac{3}{2}}}{(\mu_{xn})^{\frac{1}{2}}} \right] \left[SRM(X) + (1 + \gamma) SR^{3/2}[X]' SR^{3/2}[X] \right] \text{Diag} \left[\frac{|\sigma_{xn}|^{\frac{3}{2}}}{(\mu_{xn})^{\frac{1}{2}}} \right],$$

where $SR^{3/2}[X]$ is a $1 \times N$ vector

$$SR^{3/2}[X] = \left[(SR[X_1])^{3/2}, \dots, (SR[X_N])^{3/2} \right] = \left[\frac{(\mu_{x1})^{\frac{3}{2}}}{|\sigma_{x1}|^{\frac{3}{2}}}, \dots, \frac{(\mu_{xN})^{\frac{3}{2}}}{|\sigma_{xN}|^{\frac{3}{2}}} \right].$$

In the limit where assets have similar SRs (i.e., $SR^{3/2}[X] \sim \mathbb{1}_{1 \times N}$), the relationship simplifies to

$$(B.6) \quad \Sigma_x^{RP}(\gamma) \sim SRM(X) + (1 + \gamma) \mathbb{1}_{N \times 1} \mathbb{1}_{1 \times N}.$$

Relationships between RP-PCA and ISR-PCA covariance matrices: Using relation (10), we can express the original covariance matrix Σ_x in terms of the ISR-PCA covariance matrix $ISRM(X)$,

$$\Sigma_x = \text{Diag} [\mu_{xn}] ISRM(X) \text{Diag} [\mu_{xn}].$$

Substituting the above expression of Σ_x into the RP-PCA covariance matrix (13) yields the the relationship between RP-PCA and ISR-PCA covariance matrices,

$$(B.7) \quad \Sigma_x^{RP}(\gamma) = \text{Diag} [\mu_{xn}] [ISRM(X) + (1 + \gamma) \mathbb{1}_{N \times 1} \mathbb{1}_{1 \times N}] \text{Diag} [\mu_{xn}].$$

In the limit where assets have mean returns (i.e., $\mu_{nx} \approx \mu_x, \forall n$), the relationship simplifies to

$$(B.8) \quad \Sigma_x^{RP}(\gamma) \sim ISRM(X) + (1 + \gamma) \mathbb{1}_{N \times 1} \mathbb{1}_{1 \times N}.$$

B.2 Currency Denomination and Moment Fusing

The construction and pricing performance of factors in the international financial market do not vary only with return transformations, but also with the numeraire currencies. Changes in currency denominations and asset return bases share a basic feature. Namely, the risk space (concerning 2nd moment of returns) is identical for asset returns denominated in different nu-

numeraire currencies, just like it is identical for original and transformed asset returns. However, these changes differ in another crucial feature. That is, the return space (concerning both 1st and 2nd moments of returns) varies with the currency denomination (while it stays unchanged under return transformations) because the risk return tradeoff and compensation are specific to the home-country perspective. Furthermore, interest rate differentials present well-known proxies for currency risk premia. These features differentiate and enrich the moment-fusing constructions of FX pricing factors as the current section elaborates.

We assume an integrated, frictionless and arbitrage-free international financial (FX) market of $I + 1$ (one home and I foreign) countries, $i \in \mathcal{I} \equiv \{1, \dots, I + 1\}$. Adopting a generic base (i.e., home or numeraire) currency H , we consider a basic long-short strategy of borrowing the home currency H and lending a foreign currency $i \in \mathcal{I} \setminus H$, for the period from t to $t + 1$. The realized return at $t + 1$ in the home currency of this is

$$(B.9) \quad CT_{i/H,t+1} = \frac{S_{i/H,t+1}}{S_{i/H,t}}(1 + r_{it}) - (1 + r_{Ht}), \quad i \in \mathcal{I} \setminus H$$

where $r_{i,t}$ is country i 's risk-free rate and $S_{i/H,t}$ is the exchange rate of currency H per unit of currency i at t . Several observations are in order. First, recall that because returns of long-short strategies (B.9) are excess returns, any of their linear combinations are also excess returns (of some long-short portfolios). As these basic strategies span the FX market, all (more sophisticated) currency strategies are portfolios of these basic strategies, e.g., the generic $CT_{L/B}^H$ that borrows and lends respectively currency B and L , seen below (26). Alternatively, we can also express currency returns in terms of forward exchange rates using the covered interest rate parity (CIP). CIP is a no-arbitrage equality between the interest rate differential and forward exchange rate discount, $r_{i,t} - r_{H,t} = \frac{S_{i/H,t}}{F_{i/H,t}} - 1$, where $F_{i/H,t}$ is the one-period forward exchange rate between currencies i and H . Substituting this CIP equation into the currency strategy's return $CT_{L/B,t+1}^H = \frac{S_{L/H,t+1}}{S_{L/H,t}}(1 + r_{Lt}) - \frac{S_{B/H,t+1}}{S_{B/H,t}}(1 + r_{Bt})$ transforms it into

$$(B.10) \quad CT_{L/B,t+1}^H = \ln \frac{S_{L/H,t+1}}{S_{B/H,t+1}} - \ln \frac{F_{L/H,t}}{F_{B/H,t}}.$$

If we further express the forward exchange rate discount $\frac{S_{i/H,t}}{F_{i/H,t}} - 1$ approximately by logarithm

quantities, $s = \ln S$, $f = \ln F$, the CIP equation becomes $r_{i,t} - r_{H,t} = s_{i/H,t} - f_{i/H,t}$. Similarly, expressing the currency strategy's return in H (borrowing B and lending L) approximately by logarithm quantities, $CT_{L/B,t+1}^H = (s_{L/H,t+1} - s_{L/H,t} - r_{Lt}) - (s_{B/H,t+1} - s_{B/H,t} - r_{Bt})$. Combining these two expressions delivers the currency return (26) in Section 5.2.

Second, conditional on the information available at time t , innovations to the future currency return $CT_{i/H,t+1}$ (B.9) arise exclusively from future exchange rate innovations in $\frac{S_{i/H,t+1}}{S_{i/H,t}}$. Therefore, the principal components of exchange rate innovations can be obtained by implementing PCA on basic currency returns. Third, PC factors obtained from the set of basic currency returns $\{CT_{i/H,t+1}\}_{i \in \mathcal{I}}$ vary with currency denomination. As all the exchange rates $\left\{\frac{S_{i/H,t+1}}{S_{i/H,t}}\right\}_{i \in \mathcal{I}}$ are quotes against the home currency H , both the mean (first moment) and volatility (second moment) components of currency returns depend on the choice of numeraire currency H .⁸ These observations imply that the base currency is essential for moment-fusing factor constructions and their pricing performance as documented in Section 5.

To quantify these observations, we employ the relationship between exchange rates and the stochastic discount factors (SDF) associated with the involved currencies. Recall that in the absence of arbitrage opportunities, the pricing in each currency i is characterized by the respective country i 's SDF M_{it} . Let this SDF process be as follows

$$(B.11) \quad M_{it+1} = 1 - r_{it} - \eta'_{i\varepsilon} \varepsilon_{t+1}, \quad i \in \mathcal{I} = \{1, \dots, I+1\},$$

where vector ε_{t+1} contains uncorrelated standard normal shocks representing various risks in the international financial market, vector $\eta_{i\varepsilon}$ contains the currency i 's prices of these risks. We further assume that the market is complete, as a result of which the exchange rate growth between currency i and the base currency H is given by the ratio of the associated SDFs

$$(B.12) \quad \frac{S_{i/H,t+1}}{S_{i/H,t}} = \frac{M_{i,t+1}}{M_{H,t+1}} = 1 - (r_{it} - r_{Ht}) - (\eta_{i\varepsilon} - \eta_{H\varepsilon})' \eta_{H\varepsilon} - (\eta_{i\varepsilon} - \eta_{H\varepsilon})' \varepsilon_{t+1}, \quad \forall i \in \mathcal{I} \setminus H,$$

⁸This is an over-representation of the base currency quantified in (B.14) below.

Using this exchange rate growth, we can express the basic currency return (B.9) as follows,

$$(B.13) \quad CT_{i/H,t+1} = -(\eta_{i\varepsilon} - \eta_{H\varepsilon})' \eta_{H\varepsilon} - (\eta_{i\varepsilon} - \eta_{H\varepsilon})' \varepsilon_{t+1}$$

The currency return (B.13) and exchange rate growth (B.12) indeed have an identical volatility component $(\eta_{i\varepsilon} - \eta_{H\varepsilon})' \varepsilon_{t+1}$. The covariance-based factor analysis then can be implemented using the set of basic long-short currency returns $\{CT_{i/H,t+1}\}_{i \in \mathcal{I}}$ by diagonalizing their $I \times I$ covariance matrix

$$(B.14) \quad \Sigma_{CT} = \begin{bmatrix} (\eta_{1\varepsilon} - \eta_{H\varepsilon})' (\eta_{1\varepsilon} - \eta_{H\varepsilon}) & \dots & (\eta_{1\varepsilon} - \eta_{H\varepsilon})' (\eta_{I\varepsilon} - \eta_{H\varepsilon}) \\ \vdots & \ddots & \vdots \\ (\eta_{I\varepsilon} - \eta_{H\varepsilon})' (\eta_{1\varepsilon} - \eta_{H\varepsilon}) & \dots & (\eta_{I\varepsilon} - \eta_{H\varepsilon})' (\eta_{I\varepsilon} - \eta_{H\varepsilon}) \end{bmatrix}.$$

Evidently, the price of risk $\eta_{H\varepsilon}$ of the denomination currency is present in every return (B.13) (every component of their covariance matrix Σ_{CT}). This common presence signifies the over-representation and influence of the base currency in determining factors from the covariances of currency returns. The diagonalization of the above covariance matrix, $W_H' \Sigma_{CT} W_H = \text{Diag} [\lambda_{i/H}]$, yields the FX PC factors associated with the base currency,

$$(B.15) \quad \underbrace{[\Pi_{1/H} \dots \Pi_{I/H}]}_{T \times I} = \underbrace{[CT_{1/H} \dots CT_{I/H}]}_{T \times I} \underbrace{W_H}_{I \times I},$$

where $T \times 1$ column $\Pi_{i/H}$ denotes the i -th FX principal factor in the base currency. As these PCs are traded portfolios of basic currency strategies (B.13) denominated in the base currency H with weights W_{Hij} , their first (risk premia $\{\bar{\Pi}_{i/H,t}\}$) and second moments (risk loadings $\{\eta_{i/H}\}$) are linear combinations of currency return moments with same weights,

$$(B.16) \quad \Pi_{i/H,t+1} \equiv \bar{\Pi}_{i/H,t} + \eta'_{i/H} \varepsilon_{t+1}, \quad i \in \mathcal{I},$$

$$\text{with: } \begin{cases} [\eta_{1/H}, \dots, \eta_{I/H}] = -[(\eta_{1\varepsilon} - \eta_{H\varepsilon}), \dots, (\eta_{I\varepsilon} - \eta_{H\varepsilon})] W_H \\ [\bar{\Pi}_{1/H,t}, \dots, \bar{\Pi}_{I/H,t}] = -[(\eta_{1\varepsilon} - \eta_{H\varepsilon})' \eta_{H\varepsilon}, \dots, (\eta_{I\varepsilon} - \eta_{H\varepsilon})' \eta_{H\varepsilon}] W_H \end{cases}$$

The similarity between PCs' construction (B.15) and their risk loadings (B.16), each being obtained by the rotation W_H from the currency returns $\{CT_{i/H}\}$ and their volatilities $\{\eta_{i\varepsilon} - \eta_{H\varepsilon}\}$, allows us to visualize the variation of PCs $\{\Pi_{i/H}\}$ (which are $T \times 1$ vectors in the return space) and their risk loadings $\{\eta_{i/H}\}$ (which are $I \times 1$ vectors in the risk space) with the currency denomination. Figure B.1 illustrates this variation of the PC factors with the base currency in the risk space. In a base currency A , the standard PCA identifies PCs as orthogonalized factors explain-

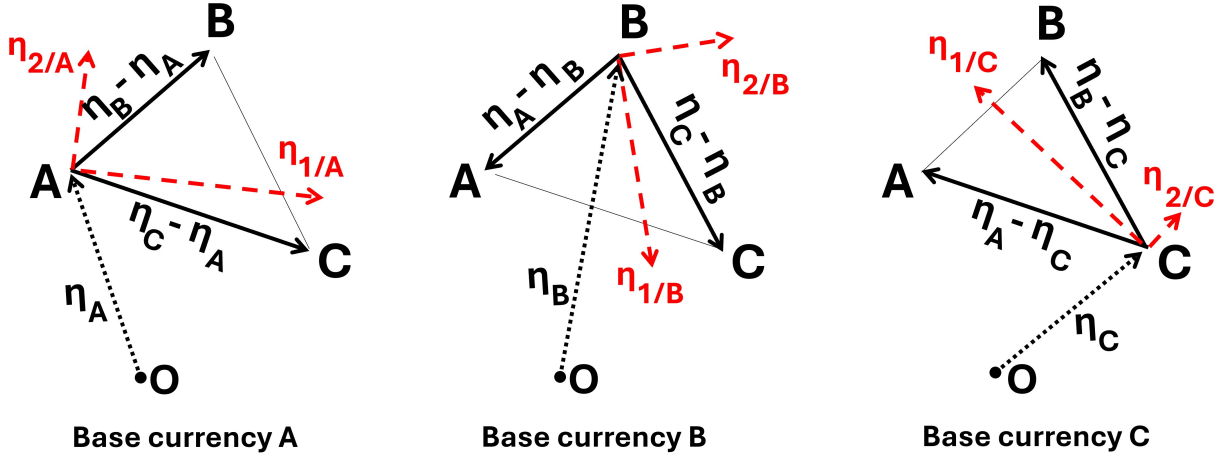


Figure B.1: This diagram illustrates principal components associated with different base currencies in the risk space. The i -th principal component in the base currency H is associated with and uniquely characterized by its risk loading vector $\eta_{i/H}$ (red arrow) in the risk space (Equation (B.16)). The base currency respectively is A (left panel), B (middle panel), and C (right panel). In a (any) base currency H : (i) per Equation (B.16), the determination of PCs involves only the relative prices of risks $\{\eta_i - \eta_H\}$, $i \neq H$, $H \in \{A, B, C\}$, and (ii) per Equation (B.16), the pricing of the PCs' risks further involves the absolute price of risk η_H in the base currency, where absolute η_H is measured from the coordinate origin O .

ing the covariance structure of risk loading vectors $\{\eta_{i\varepsilon} - \eta_{A\varepsilon}\}_{i \in \mathcal{I} \setminus A}$ (B.16), all of which clearly are characterized relative to the base country's price of risk vector $\eta_{A\varepsilon}$. A change from the base currency A to another base currency B therefore amounts to changing the reference point from A to B (the triangular ABC in Figure B.1). First, such a change of the reference point can dramatically change the PC system, both in PCs' directions (i.e., risks represented by PCs) and their magnitudes (i.e., shares of covariations explained by PCs). This indicates a significant variation of PCs' and their statistical properties with the base currency. Second, PCs' mean excess returns $\{(\eta_{i\varepsilon} - \eta_{A\varepsilon})' \eta_{A\varepsilon}\}_{i \in \mathcal{I} \setminus A}$ (B.16), which quantify how the risks represented by PCs are priced in the base currency, also involves the price of risk $\eta_{A\varepsilon}$ of the base currency (the quadrilateral $ABCO$).

The exogeneity of the origin O from A, B, C allows for significant flexibility and divergences of PCs' pricing properties across different base currencies. Moment fusing employs interest differentials as a well-known proxy for the risk premia of currency strategies and the above variations to construct denomination-specific factors and pricing models. Our empirical analysis of Section 5 demonstrates significant but also variable pricing performance of moment-fusing factors in different currency denominations.

B.3 Pricing Measures

For a self-contained connection between actual test statistics and the conceptual description of the moment-fusing approach, 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, $\tilde{\Pi} = \tilde{X}W$), 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 $X_{tn} = \alpha_n + \sum_{k=1}^K \Pi_{tk} \beta_{kn} + \varepsilon_{tn}$, $n \in \{1, \dots, N\}$, or in matrix form (stacking asset returns into columns),

$$(B.17) \quad X_{T \times N} = \mathbb{1}_{T \times 1} \alpha_{1 \times N} + \Pi_{T \times K} \beta_{K \times N} + \varepsilon_{T \times N}.$$

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

$\tilde{M}_{K\parallel} = -\tilde{\Pi} \left[\frac{1}{T} \tilde{\Pi}' \tilde{\Pi} \right]^{-1} \mu'_{\Pi}$, where $1 \times K$ vector μ_{Π} contains the means (or risk premia) of K factors.⁹

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-

⁹Recall 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.

Shanken (GRS) test statistic, and (iv) share of unexplained (idiosyncratic) return variation.

Maximum attainable Sharpe ratio (max SR): The max SR of the model G_K , denoted hereafter as $\max SR(G_K)$, 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, $\max SR(G_K) = \sqrt{\text{Var}[M_{K\parallel}]}$. 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

$$(B.18) \quad [\max SR(G_K)]^2 = \sum_{n=1}^K (SR[\Pi_n])^2, \quad \text{with} \quad \Pi_{kt} = \sum_{n=1}^N X_{nt} W_{nk}.$$

When all factors are retained ($K = N$), Proposition 1 implies that $\max SR(G)$ 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 $\max SR(G_K)$. 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

$$(B.19) \quad \alpha_n = \mu_{xn} + \widetilde{M}'_{K\parallel} \widetilde{X}_n = \mu_{xn} - \mu'_\Pi \widetilde{\Pi} \left[\frac{1}{T} \widetilde{\Pi}' \widetilde{\Pi} \right]^{-1} \widetilde{X}_n, \quad \alpha = \mu_x - \mu'_\Pi \widetilde{\Pi} \left[\frac{1}{T} \widetilde{\Pi}' \widetilde{\Pi} \right]^{-1} \widetilde{X}$$

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

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. Giv-

en 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 a uniform leverage operation on their excess returns $Y_n = \kappa X_n$, $n \in \{1, \dots, N\}$ (3) scales linearly (hence arbitrarily, via the choice of κ), $RMSE_\alpha(Y) = \kappa RMSE_\alpha(X)$. With regard to this amenability, our empirical analysis employs 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 (B.17). The GRS null hypothesis is that all pricing errors α 's in (B.17) 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 $N \times N$ full-rank covariance matrix Σ_ε , then the GRS test statistic defined as

$$(B.20) \quad GRS = \frac{T}{N} \frac{T - N - K}{T - K - 1} \frac{\hat{\alpha} \hat{\Sigma}_\varepsilon^{-1} \hat{\alpha}'}{1 + \mu_\Pi' \hat{\Sigma}_\Pi^{-1} \mu_\Pi},$$

follows a F -distribution with N and $T - N - K$ degrees of freedom (or, $F_{N, T-N-K}$) under the null hypothesis. In (B.20), $\hat{\Sigma}_\Pi = \frac{1}{T} \tilde{\Pi}' \tilde{\Pi}$, and $\hat{\Sigma}_\varepsilon = \frac{\tilde{\varepsilon}' \tilde{\varepsilon}}{T-K-1}$, and $\hat{\varepsilon}$ and $\hat{\alpha}$ are least squares estimates from (B.17). A statistically sufficiently low value of GRS test statistic (or, a small enough p -value) is desirable.

Share 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 (B.17) 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

$$(B.21) \quad \sigma_e = 1 - \frac{\sum_{n=1}^N \sum_{k=1}^K Covar(X_n, \Pi_k \hat{\beta}_{kn})}{\sum_{n=1}^N Var(X_n)} = 1 - \frac{Trace[\tilde{X}' \tilde{\Pi} \hat{\beta}]}{Trace[\tilde{X}' \tilde{X}]},$$

where $\hat{\beta}$'s are least squares estimates for the asset loadings on factors (B.17). When factors

$\{\Pi_n\}_{n=1}^K$ are mutually uncorrelated (or factors are re-orthogonalized) such as PCs, the % of unexplained return variation becomes $\sigma_e = 1 - \frac{\sum_{k=1}^K \lambda_k}{\sum_{k=1}^N \lambda_k}$, where $\lambda_n = Var(\Pi_n)$, $n \in \{1, \dots, N\}$. 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 ($\Pi_n(\kappa_n) = \kappa_n \Pi_n$, $n \in \{1, \dots, N\}$, 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 the first moment of returns, and given the amenability of σ_e and RMSE in this process, focuses on $maxSR$ as a primary pricing measure.