Stochastic Discount Factors with Cross-Asset Spillovers *

Doron Avramov

Reichman University (IDC), Herzliya, Israel

Xin He

University of Science and Technology of China

May 20, 2025

Abstract

This paper develops a structured asset-pricing framework for estimating the stochastic discount factor (SDF) by aggregating firm-level signals while accounting for informational linkages across assets. The method identifies which characteristics are most relevant for pricing and uncovers the directional flow of predictive influence among firms. Empirically, the resulting SDF delivers strong out-of-sample performance across asset universes and market regimes. Moreover, large, low-turnover firms emerge as central nodes in the information network. The framework provides a transparent and economically grounded view of the informational structure embedded in return dynamics.

Key Words: Asset Pricing, Cross-Asset Spillover, Connection Matrix, Sharpe Ratio, Stochastic Discount Factor.

JEL classification: C1, G11, G12.

*We are grateful to Lin Will Cong, Shuyi Ge, Michael Gofman, Semyon Malamud, Zilong Niu, Alex Philipov, Gil Segal, Robert Stambaugh, Yinan Su, Jingzhou Yan, Jingyi Yao, Guofu Zhou and seminar and conference participants at Sichuan University, Southwestern University of Finance and Economics, Xiamen University, and 2025 Annual Meeting of Chinese Statistical Association of Young Scholars for constructive discussions and feedback.

Avramov (E-mail: doron.avramov@runi.ac.il) is at Reichman University (IDC), Herzliya, Israel.

He (E-mail: xin.he@ustc.edu.cn) is at International Institute of Finance, Faculty of Business for Science and Technology, School of Management, University of Science and Technology of China, Hefei, China.

1 Introduction

A defining objective of empirical asset pricing is to identify firm-level signals that help explain the cross-section of expected stock returns, whether due to persistent mispricing or exposures to risk factors. The prevailing approach, grounded in the assumption of self-predictability, posits that a firm's own characteristics forecast its own return (e.g., Cochrane, 2011; Harvey et al., 2016). A complementary line of research, however, underscores the importance of cross-predictability—the idea that the returns or characteristics of one asset can help forecast the returns of others (e.g., Lo and MacKinlay, 1990; Hou, 2007; Cohen and Frazzini, 2008; Cohen and Lou, 2012; Kelly et al., 2023). In particular, lead–lag effects—whereby price movements or information in one firm precede and help predict subsequent movements in related firms—provide an intuitive mechanism for these spillovers: through staggered information diffusion, industry peer influence, or supply-chain linkages, shocks propagate gradually across assets. Exploiting these spillovers can materially enhance forecasting performance beyond what self-predictive signals alone deliver.

While recent advances mark substantial progress, several foundational questions remain unresolved. Chief among them: how should a mean–variance optimizing investor integrate multiple predictive signals when returns exhibit cross-asset dependencies? And how can one jointly capture both the relevance of signals and the structure of inter-asset predictability in a way that enhances portfolio efficiency while maintaining interpretability?

This paper proposes a systematic framework to address these challenges by combining firm-level signals via a flexible weighting vector and modeling predictive spillovers through a structured connection matrix. The resulting strategy is derived by maximizing the Sharpe ratio and is solved analytically using a tractable generalized eigenvalue decomposition. This formulation naturally links the optimal strategy to the stochastic discount factor (SDF; e.g., Cochrane (2009); Back (2017)), which, in this context, takes the form of a single factor that prices the cross-section of asset returns.

While the generalized eigenvalue problem provides a population-level characterization of the Sharpe-optimal SDF, our empirical implementation relies on a regression-based procedure better suited for high-dimensional settings. Specifically, we build on the approach of Britten-Jones (1999) and apply ridge-type regularization with a single hyperparameter to estimate the signal weights and the connection structure. This method converges to the eigenvalue solution in large samples

while offering improved numerical stability and interpretability. In contrast to expected-return maximization—which under certain conditions, such as unconstrained portfolios or L_1 penalties, places all weight on a single dominant predictor—Sharpe ratio maximization yields a diversified weighting scheme across signals, enhancing robustness. Our framework complements recent advances in SDF estimation using firm characteristics (e.g., Kelly et al. (2019); Kozak et al. (2020); Lettau and Pelger (2020); Gu et al. (2021); Chen et al. (2024); Feng et al. (2024); Didisheim et al. (2024); Cong et al. (2025); Liu et al. (2025)) by explicitly modeling inter-asset return linkages in a structured and interpretable manner.

To build intuition, we begin with a low-dimensional toy example based on five well-known firm characteristics and nine portfolios sorted by size and book-to-market. This simplified setting allows us to present the estimated signal weights, cross-asset linkages, and resulting trading strategy in full detail. Performance is evaluated using a rolling out-of-sample procedure, where the strategy is re-estimated each month using the prior 10-year data. Even in this controlled environment, the Sharpe-maximizing strategy based on cross-stock predictability achieves an annualized Sharpe ratio of 1.38, compared to 1.02 under the self-predictive benchmark. The improvement reflects not only return spillovers captured by the off-diagonal elements of the connection matrix but also meaningful shifts in signal relevance. These results suggest that incorporating cross-predictability reshapes portfolio construction in fundamental ways.

We then scale the framework to a comprehensive empirical setting using 138 firm-level signals from the Jensen et al. (2023) dataset. Our primary investment universe consists of 138 univariate spread portfolios spanning 1960–2023. We also consider a broader set of 544 bi-variate portfolios sorted by firm size and a secondary characteristic. Applying the same rolling 10-year estimation scheme, the Sharpe-maximizing strategy attains annualized Sharpe ratios of 2.19 and 3.62 on the spread and bi-sort portfolios, respectively—consistently outperforming both self-predictive benchmarks and expected-return strategies.

To assess robustness, we evaluate performance across different market environments. We split the sample by investor sentiment Baker and Wurgler (2006) and by volatility regimes based on the VIX index. The Sharpe-maximizing strategy maintains strong performance across all subsamples. For example, in high-sentiment periods, the strategy delivers a Sharpe ratio of 2.15 on spread port-

folios and 3.96 on bi-sort portfolios. Even in low-sentiment or high-volatility regimes—conditions that typically challenge many anomaly-based strategies—the strategy sustains Sharpe ratios above 2.0. These results contrast with the more state-dependent performance of return-maximizing portfolios as well as individual anomalies, which perform well only in certain regimes.

The Sharpe ratio–maximizing strategy, or the stochastic discount factor (SDF), defines a single factor that ex ante prices the cross-sectional variation in expected returns of test assets. We examine whether this factor's returns are priced by leading asset pricing models and find sizable, statistically significant alphas relative to a broad set of benchmarks. These include the liquidity factor (Pastor and Stambaugh, 2003), the Fama–French five-factor model (Fama and French, 2015), the q-factors (Hou et al., 2015), the mispricing factors (Stambaugh and Yuan, 2017), the behavioral factors (Daniel et al., 2020), and a comprehensive fourteen-factor model. Across all specifications, the strategy's alphas remain both statistically and economically significant—approximately 0.16% per month, with t-statistics exceeding 11—suggesting that existing models do not fully capture return variation because they overlook the predictive content embedded in cross-asset spillovers.

Upon optimizing the Sharpe ratio, we offer insight into the underlying economic drivers of return predictability. By examining the estimated signal weights —which determine how firm-level characteristics contribute to the Sharpe-maximizing stochastic discount factor—we identify the signals that play a central role in shaping the strategy. These weights reveal that the most influential predictors are concentrated in the *investment*, *value*, and *profitability* categories, with top signals such as *liquidity of book assets*, *dividend yield*, and *return on equity* consistently receiving high importance. In contrast, return-based signals such as momentum, short-term reversal, and seasonality exhibit persistently low weights. This pattern suggests that the cross-predictive SDF is anchored in stable firm fundamentals rather than transitory market signals.

In optimizing the Sharpe ratio, we also obtain a connection matrix that encodes the predictive relationships across stocks. Each entry $\Psi_{i,j}$ reflects the extent to which signals from asset i forecast the returns of asset j, while diagonal elements represent self-predictive strength. Empirically, the average off-diagonal entry is substantial—often exceeding the average diagonal—indicating that inter-asset predictive linkages carry more information than self-predictive signals alone. Aggregating rows and columns of the matrix following Diebold and Yılmaz (2014), we uncover a directional

structure: certain stocks consistently act as net transmitters of predictive signals, while others serve primarily as receivers. Transmitters are typically large and low-turnover, whereas receivers tend to be smaller, high-turnover stocks with characteristics such as value orientation, high profitability, low investment activity, and strong past returns.

Finally, it is worth noting that the Sharpe ratio of the cross-predictive strategy is time-varying and has declined notably since 2000. In the 1990s, the strategy delivered exceptional performance, with Sharpe ratios exceeding 3 on spread portfolios and above 5 on bi-sort portfolios. However, performance attenuated in the post-2000 period, mirroring the broader decline in self-predictability. For instance, Green et al. (2017) document that many return anomalies became less profitable after 2003, attributing the decline to the widespread adoption of anomaly-based strategies, improved market liquidity, and the growth of passive ETF investing.

Despite this attenuation, our strategy maintained strong performance from 2000 to 2023, achieving Sharpe ratios of 1.33 (spread portfolios) and 2.42 (bi-sort portfolios)—substantially higher than those of standard benchmark factors: 0.41 (market), 0.20 (size), 0.20 (value), 0.54 (profitability), 0.43 (investment), and 0.09 (momentum). By the end of 2023, five-year trailing Sharpe ratios declined to approximately 1.5 for the bi-sort strategy and 1.0 for the spread strategy, yet both remained consistently superior to traditional factors even in recent years.

Two important works by Kelly et al. (2023) and He et al. (2024) are closely related to ours. The former develops principal portfolios that maximize expected returns using a single predictor, while the latter extends this framework to accommodate multiple predictors. Our study introduces a structured and interpretable analytical framework designed to estimate the stochastic discount factor (SDF) in the presence of cross-asset linkages. The estimation procedure relies on a single hyperparameter, and empirical evidence shows that the resulting SDF consistently outperforms both self-predictive models and expected-return maximizers that incorporate cross-predictability.

The paper proceeds as follows. Section 2 presents the econometric framework for making investment decisions based on cross-predictability and multiple signals. Section 3 outlines the estimation methodology. Section 4 describes the data. Section 5 reports the empirical findings. Section 6 concludes.

2 Econometric Framework

We consider an investment universe of N risky assets. At time t, a hypothetical investor observes a signal matrix $S_t \in \mathbb{R}^{N\times M}$, whose ith row collects the M predictive attributes (e.g., market capitalization, valuation ratios, profitability, investment, past returns) for asset i. Each column of S_t is standardized to have zero cross-sectional mean and unit variance. The vector of excess returns at any future date $s > t$ is denoted by $r_s \in \mathbb{R}^N$. We begin by analyzing linear portfolio strategies that map the signal matrix into asset weights, and then proceed to consider nonlinear strategies.

2.1 Linear Strategy

A linear trading strategy with multiple signals and cross-predictability is formulated as

$$\omega_t' = \Lambda' \, S_t' \, \Psi, \tag{1}$$

where $\omega_t \in \mathbb{R}^N$ is the vector of portfolio weights, $\Lambda \in \mathbb{R}^M$ assigns loadings to each signal, and $\Psi \in \mathbb{R}^{N\times N}$ captures how signals for one asset influence positions not only in that asset but also in all others. In particular, the weight on asset i is obtained by multiplying Λ', S_t', and the ith column of Ψ, so that all the signals encapsulated in S_t contribute to the position in each asset.

We construct managed portfolio returns in excess of the risk-free rate by interacting future returns with the current values of predictive signals:

$$\Pi_s = \big(I_N \otimes r_s\big) \, S_t. \tag{2}$$

In this expression, Π_s is an $N^2 \times M$ matrix of managed portfolio returns, I_N is the $N \times N$ identity matrix, and $\otimes$ denotes the Kronecker product.

The expected returns on these managed portfolios are then defined as $\Pi = E\big[\Pi_s\big]$. Additionally, define

$$\Phi = \mathrm{vec}\big(\Psi'\big), \tag{3}$$

so that $\Phi \in \mathbb{R}^{N^2}$. The vectorized Φ and the matrices Π_s and Π streamline later expressions for portfolio outcomes.

To aid interpretation, limit extreme positions, and stabilize estimation, we impose Euclidean-norm constraints:

$$\Lambda'\Lambda = 1, \quad \Phi'\Phi = 1. \tag{4}$$

From a Bayesian perspective, these constraints correspond to zero-mean Gaussian priors on Λ and Φ, yielding ridge-type regularization that penalizes large parameter values.

Proposition 1 formulates the realized return of the strategy in a convenient form, along with the expected return and Sharpe ratio. Appendix A provides the proof.

Proposition 1. *The investment metrics are as follows:*

- *The realized and expected returns can be expressed as*

$$\pi_s = \Lambda'\Pi_s\Phi, \tag{5}$$

$$E(\pi_s) = \Lambda'\Pi\Phi. \tag{6}$$

- *The square of the Sharpe Ratio (SR^2) is given by the following two equivalent expressions:*

$$SR^2 = \frac{\Lambda' A_\Phi \Lambda}{\Lambda' B_\Phi \Lambda}, \tag{7}$$

$$SR^2 = \frac{\Phi' A_\Lambda \Phi}{\Phi' B_\Lambda \Phi}. \tag{8}$$

Here, $A_\Phi = \Pi'\Phi\Phi'\Pi$, $B_\Phi = (\Phi' \otimes I_M)\Sigma_\Phi(\Phi \otimes I_M)$, Σ_Φ is the covariance matrix of vec(Π_s'), and I_M is the identity matrix of order M. Similarly, $A_\Lambda = \Pi\Lambda\Lambda'\Pi'$, $B_\Lambda = (\Lambda' \otimes I_{N^2})\,\Sigma_\Lambda\,(\Lambda \otimes I_{N^2})$, Σ_Λ is the covariance matrix of vec(Π_s), and I_{N^2} is the identity matrix of order N^2.

We make several notes regarding Proposition 1.

First, it extensively utilizes the vectorized Φ, which retains all cross-predictive relationships among assets encapsulated in Ψ. Hence, the information for cross-predictability is fully preserved in Φ, ensuring that the trading strategy remains grounded in the same foundational predictive content.

Second, we consider two investors optimizing distinct objectives. Investor I maximizes expected return, focusing solely on returns without explicit risk considerations, while Investor II

maximizes the Sharpe ratio, balancing return and risk. Although both rely on the same expressions for expected return and the Sharpe ratio, their optimal estimates of Λ and Φ differ: maximizing expected return reduces to optimizing a bilinear form with closed-form solutions, whereas maximizing the squared Sharpe ratio requires solving a generalized eigenvalue problem via an iterative algorithm.

Notably, to maximize the squared Sharpe ratio, it is essential to employ both representations of the Sharpe ratio in Proposition 1 when estimating the optimal values of Λ and Φ, with explicit solutions to the corresponding maximization problems—both for expected return and for Sharpe ratio—provided later.

Third, the expression for investment return offers an intuitive economic interpretation of our trading strategy. Recall that Π denotes the matrix of managed-portfolio expected returns, with each of its N^2 rows representing the expected value of one asset's return multiplied by one of the M signals across the N assets. Under the normalization $\mathbb{E}[S_t] = 0$, Π simplifies to the covariance matrix between future asset returns and contemporaneous signal values. If characteristic m of stock j helps predict the future return of stock i, the corresponding element of Π will be nonzero, reflecting this predictability.

Thus, in this framework, Λ assigns relative weights to signals, Φ encodes inter-asset interactions, and together they operate on the predictive matrix Π to optimize investment metrics.

Fourth, the expected return of the trading strategy can alternatively be expressed as

$$E(\pi_s) = \sum_{m=1}^{M} \Lambda_m \mu_m, \tag{9}$$

where $\mu_m = \sum_{p=1}^{N^2} \Pi_{pm} \Phi_p$ represents a weighted combination of portfolio expected returns, with Π_{pm} denoting the expected return of the corresponding managed portfolio and Φ_p capturing the strength of the p-th relationship within the strategy.

This expected-return expression is informative because it demonstrates that, whether subject to an L_1 constraint or left unconstrained, the optimal solution is a corner solution: the trading strategy is entirely driven by the predictor with the largest absolute value of μ_m, denoted predictor j, with $|\Lambda_j| = 1$ and all other elements of Λ equal to zero. In contrast, under an L_2 constraint, the optimal

Λ (given Φ) is proportional to the M-vector that collects the μ_m values. By comparison, Sharpe ratio maximization effectively harnesses the benefits of diversification across predictors, assigning meaningful weight to multiple signals.

Fifth, the realized return π_s of the maximum Sharpe ratio portfolio is proportional to the stochastic discount factor (SDF). This follows directly from the fundamental asset pricing representation (Cochrane, 2009; Back, 2017):

$$M_s = 1 - b^\top r_s, \quad \text{with} \quad \mathbb{E}[M_s, r_s] = 0, \tag{10}$$

where M_s denotes the pricing kernel and b is a vector of slope coefficients. Solving this system reveals that b corresponds to the tangency portfolio weights. Consequently, projecting r_s onto π_s—the realized return of the maximum Sharpe ratio strategy—yields a population regression with zero intercept, where the slope coefficients represent the SDF loadings. This allows us to estimate the SDF in the presence of cross-stock return linkages—a capability that, to our knowledge, is novel.

For comparison, Kelly et al. (2023) decompose the returns of linear strategies with cross-stock predictability into alpha and beta components when maximizing expected return. In their setting, alpha arises because the optimal strategy is mean–variance efficient, even when expected return is maximized under a volatility constraint. In our empirical analysis, we show that the maximum expected return and maximum Sharpe ratio strategies—both accounting for cross-asset spillovers—differ substantially, with the latter delivering significantly higher Sharpe ratios across the full sample and in both expanding and contracting regimes.

2.2 Zero-Cost and Leverage Constraints

Up to this point, we have not explicitly imposed any constraints on the strategy's positions. The empirical asset pricing literature typically requires that a trading strategy, factor, or anomaly take the form of a long-short portfolio. In other words, the total cost must be zero, and the total leverage must equal two.

The proposition below imposes zero-cost constraints on the strategy.

Proposition 2. *A zero-cost trading strategy can be expressed as follows:*

$$\omega_t' = \Lambda' S_t' \Psi - \frac{1}{N} \Lambda' S_t' \Psi A, \tag{11}$$

$$= \Lambda' S_t' \Psi \Theta, \tag{12}$$

where A is an $N \times N$ matrix, with each element set to one, and $\Theta = I_N - \frac{1}{N} A$.

Notice that $\omega_t' \iota_N = 0$, where ι_N is an N-vector of ones. Fortunately, all previous derivations remain valid under the zero-cost constraint.

The necessary modifications are as follows. Define $\Pi_{si} = \Theta(r_s S_{it}')$ for each $i = 1, 2, \ldots, N$, and construct Π_s by vertically stacking Π_{si}. All investment metrics in Proposition 1 can then be re-derived under the zero-cost constraint.

In Appendix B, we demonstrate that the zero-cost constraint reduces the expected profitability of the trading strategy. However, this constraint is essential for ensuring comparability across strategies.

In our empirical analyses, we primarily focus on zero-cost strategies, where the long and short positions are of equal magnitude by construction. To further ensure comparability, we rescale these positions so that total portfolio leverage equals two. This adjustment aligns our strategies with standard practice in the literature (e.g., Fama and French, 1993).

2.3 Economic Restrictions

The decision variable Φ consists of N^2 elements, capturing the network among investable assets. However, as N grows large, the computational burden increases exponentially, posing challenges for both estimation and out-of-sample (OOS) performance. Thus, prior economic knowledge about asset linkages can be incorporated into Φ, allowing for a more structured and interpretable model. The use of economic restrictions has proven beneficial in high-dimensional modeling and machine learning applications in finance; see, e.g., Avramov, Cheng, and Metzker (2023).

In our context, if cross-predictability primarily flows from big stocks to small stocks but not vice versa, it is reasonable to impose restrictions on Φ by setting to zero the elements that represent small stock signals predicting large stock returns. This constraint aligns with the findings of Lo and

MacKinlay (1990), who suggest that big stocks tend to lead small stocks.

By refining Φ to emphasize meaningful economic connections, one can enhance both the interpretability and predictive performance of the strategy. This approach can be applied to a broad range of economically meaningful linkages, such as common ownership, industry classification, liquidity spillovers, and supply chain relationships, among others.

3 Estimating the Unknown Parameters

We provide methods for estimating the unknown parameters underlying the trading strategy by optimizing investment metrics.

3.1 Maximizing Expected Return

Proposition 3 presents the solution for the strategy that maximizes expected return. For notational clarity, we focus on the linear case; however, the results extend to nonlinear strategies by expanding the set of signals S_t to include polynomial transformations and random Fourier features.

Proposition 3. *By the Singular Value Decomposition (SVD),* Π *can be decomposed as*

$$\Pi = U \Lambda_{\Pi} V', \tag{13}$$

where U *is an* $N^2 \times N^2$ *orthogonal matrix,* Λ_{Π} *is an* $N^2 \times M$ *diagonal matrix of singular values, and* V' *is an* $M \times M$ *orthogonal matrix.*

The estimated parameters that maximize expected returns are given by

$$\hat{\Lambda} = V(:, 1), \tag{14}$$

$$\hat{\Phi} = U(:, 1). \tag{15}$$

The estimated values of Λ and Φ are selected as the first singular vectors from the matrices V and U, respectively. This ensures that the optimal trading strategy formulation leverages the directions that capture the maximum variance, without implying dimension reduction. Instead, we

utilize the leading singular vectors for their high explanatory power, capturing the most significant patterns in the data.

3.2 Maximizing Sharpe Ratio

Propositions 4 and 5 provide the solution and the estimation for the strategy to maximize squared Sharpe ratio. Appendix C provides the proof and detailed derivations.

Proposition 4. *Assume that Φ is given. Based on (7), define*

$$C_\Phi = B_\Phi^{-1} A_\Phi. \tag{16}$$

The optimal Λ is the principal eigenvector $\Lambda_{\max}$ of the eigenvalue problem

$$C_\Phi \, \Lambda = \lambda \, \Lambda. \tag{17}$$

Similarly, assume Λ is given. Based on Equation (8), define $C_\Lambda = B_\Lambda^{-1} A_\Lambda$. The optimal value for Φ is the largest eigenvector Φ_{max} of the following eigenvalue problem:

$$C_\Lambda \Phi = \lambda \Phi. \tag{18}$$

The optimal solutions for Λ and Φ are obtained by iteratively applying these two equations until convergence. We further rescale each solution to have unit norm.

In this way, we utilize both alternative expressions for the Sharpe ratio in Proposition 1 to iteratively estimate the optimal parameters Λ and Φ. However, the eigenvalue problems in (17) and (18) require computing the inverse of large matrices, which is challenging in high-dimensional settings. To address this, we propose the following proposition to iteratively estimate Λ and Ψ.

Proposition 5. *Consider a set of managed portfolios χ_Φ of dimension $T \times M$:*

$$\chi_\Phi = \Pi' \Phi. \tag{19}$$

The problem in (17) is essentially an asset-allocation exercise: it seeks to maximize the squared Sharpe ratio

by investing in χ_Φ composed of M assets. This is equivalent to estimating Λ as the mean-variance efficient portfolio weights.

Following Britten-Jones (1999), the estimate of Λ is obtained from the following regression:

$$\mathbf{1} = \chi_\Phi \Lambda + \mathbf{u}, \tag{20}$$

where $\mathbf{1}$ is a T-vector of ones and T denotes the sample size. To handle high-dimensional settings, we adopt ridge regression (Kelly and Xiu, 2023; Shen and Xiu, 2024; Didisheim et al., 2024). The estimator for Λ is then given by:

$$\hat{\Lambda} = (\chi_\Phi' \chi_\Phi + \lambda I_M)^{-1} \chi_\Phi' \mathbf{1}, \tag{21}$$

where λ is a Ridge-type parameter that shrinks the regression coefficients towards zero. [1]

Similarly, we define a set of managed portfolios χ_Λ of dimension $T \times N^2$:

$$\chi_\Lambda = \Pi \Lambda. \tag{22}$$

The problem in (18) is another asset allocation exercise: it seeks to maximize the squared Sharpe ratio by investing in χ_Λ. The estimator for Φ is

$$\hat{\Phi} = (\chi_\Lambda' \chi_\Lambda + \lambda I_{N^2})^{-1} \chi_\Lambda' \mathbf{1}. \tag{23}$$

We emphasize three key aspects of Sharpe-ratio maximization. First, the preceding propositions recast the problem as a managed-portfolio exercise, yielding the optimal weights (Λ^*, Φ^*) for the tangency portfolio—or equivalently the stochastic discount factor—given the model parameters. Second, we impose a common ridge penalty λ when estimating both Λ and Φ, which enforces uniform shrinkage across all components. This shared penalty simplifies exposition, promotes replication, and helps guard against overfitting in finite samples.

Third, the generalized-eigenvalue result provides a population-level characterization of the Sharpe-ratio maximizer, but in practice we replace the unknown moment matrices with their sample analogues and the same ridge penalty. Framing estimation as a single ridge-penalized regres-

[1] Following Didisheim et al. (2024); Cong et al. (2025), we choose a minimal value $\lambda = 10^{-4}$ in empirical analysis. The additional tests in Appendix D show that the OOS Sharpe ratio is robust to an array of values for λ.

sion—rather than performing an explicit generalized eigen-decomposition—allows us to recover the optimal SDF direction directly, enhances numerical stability by shrinking weights on weak or collinear signals, and avoids the computational burden of eigen-solvers. The estimated weight vector produced by this approach coincides exactly with the theoretical maximizer in finite samples.

Our method for estimating the Sharpe ratio–maximizing strategy relates to a recent literature that applies *transformer* architectures to asset pricing, leveraging multi-headed attention mechanisms to extract and aggregate predictive signals across assets. Cong et al. (2022) introduce *AlphaPortfolio*, a deep reinforcement learning framework for portfolio optimization that incorporates cross-asset attention networks (CAAN) to model interdependencies among securities. The *AIPM* framework of Kelly et al. (2024) embeds transformer architectures into the stochastic discount factor (SDF) model, demonstrating that nonlinear information sharing across assets can significantly enhance empirical performance.

While these approaches offer substantial modeling flexibility, our framework contributes a complementary, linear formulation that emphasizes transparency and interpretability. Specifically, we represent cross-asset spillovers through a connection matrix Ψ, where each element $\Psi_{i,j}$ quantifies the predictive influence of asset i's signals on asset j's returns. Although related to the linear component of AIPM, our approach differs in that AIPM models the attention matrix as a function of asset-level signals, requiring the estimation of parameters on the order of M^3, where M is the number of characteristics. In such settings, the effects of signal relevance and inter-asset connections are embedded jointly in the signal space.

Our framework disentangles these two dimensions: signal relevance is captured by the vector Λ, while cross-asset connections are modeled separately via Ψ. The resulting parameter complexity is only of order M and N^2, respectively. This structure promotes computational efficiency, facilitates starightforward replication, and provides an economically interpretable lens through which to understand both the source of predictive power and the pattern of return spillovers across assets.

4 Data

Our dataset combines monthly stock returns from the Center for Research in Security Prices (CRSP), accounting variables from Compustat, and analyst coverage and earnings forecasts from

the Institutional Brokers' Estimate System (IBES). We assume that quarterly and annual financial statements from Compustat become publicly available four months after the end of the corresponding fiscal quarter. The full sample spans January 1963 to December 2023. Out-of-sample evaluation begins in February 1973, with estimation windows based on rolling samples of the most recent 120 monthly observations.

4.1 Predictive Characteristics

We employ 138 firm-level signals across 13 characteristic themes: Accruals, Debt Issuance, Investment, Leverage, Low Risk, Momentum, Profit Growth, Profitability, Quality, Seasonality, Size, Short-Term Reversal, and Value. These signals originate from Jensen et al. (2023).[2]

4.2 Spread Portfolios

For each of the 138 signals, we sort stocks into terciles each month and compute high-minus-low factor returns. To form factor-level signals, we aggregate stock-level signals into corresponding factor portfolios. Returns and signals are value-weighted by market equity, with individual market-equity weights winsorized at the 80th percentile of NYSE capitalization, following the data providers' recommendations.

4.3 Bi-Variate Sorting on Size and Other Characteristics

We also construct bi-variate sorted portfolios to serve as alternative investment universe. First, stocks are sorted into two size groups (big vs. small) based on market equity. Independently, each signal sorts stocks into three groups (high, medium, low). Cross-classifying these sorts produces six portfolios; we retain only the high and low portfolios for each size group, resulting in four portfolios per signal. As with the spread portfolios, returns and signals are capped-value-weighted by winsorized market equity. We omit the bi-variate portfolios for the characteristic `ami_126d` due

[2]We use the "Global Stock Returns and Characteristics" dataset under "Contributed Data Forms" on WRDS: https://wrds-www.wharton.upenn.edu/pages/get-data/contributed-data-forms/global-factor-data/. Table IA.II of Jensen et al. (2023) details the signal definitions and references. Of the original 153 signals, we exclude 15 that begin after 1963 to satisfy the sample-coverage requirements of Kelly et al. (2023). We apply standard filters to retain only observations with: (i) `excntry` = "USA", (ii) CRSP `shrcd` $\in \{10, 11\}$, (iii) CRSP `exchcd` $\in \{1, 2, 3\}$, and (iv) non-missing monthly excess return (`ret_exc`) and next-month excess return (`ret_exc_lead1m`). Each characteristic is standardized to have a mean of zero and a standard deviation of one.

to missing returns in 2023. Moreover, since size already plays a role in the sorting procedure, we consider a total of $136 \times 4 = 544$ portfolios.

Thus, we consider two investment universes: one constructed from univariate sorts comprising 138 spread portfolios, and the other from bivariate sorts comprising 544 portfolios. Each portfolio is associated with a time series of returns and 138 signal observations.

5 Results

5.1 An Illustrative Toy Example

To build intuition for the proposed framework, we construct a low-dimensional toy dataset comprising five firm characteristics—market equity (ME), book-to-market ratio (BM), operating profits to lagged book equity (OP), asset growth (INV), and 12-month momentum (MOM)—and nine portfolios formed by a 3×3 sort on ME and BM (ranging from ME1×BM1 to ME3×BM3).

This simplified setup allows us to explicitly report the estimated low-dimensional parameters Λ and Ψ, as well as the weight vector ω. It also enables a comparison of key performance metrics for: (i) strategies subject to unit-norm constraints without an explicit zero-cost requirement; and (ii) zero-cost strategies with total leverage constrained to two.

We implement expected return and Sharpe ratio maximizing strategies, as formulated in the methodological section. These strategies, which target different objectives, yield notable differences in parameter estimates and performance outcomes. Table 1 summarize the monthly average returns, monthly standard deviations, and *annualized* Sharpe ratios for each strategy over the out-of-sample period from February 1973 to December 2023.

The first two rows of the table consider the case in which the zero-cost assumption is not imposed. The results show that the strategy maximizing expected return (MR Cross) delivers a high average monthly return of 1.96%, but with substantial volatility (standard deviation of 21.89%), yielding a Sharpe ratio of just 0.31.

The Sharpe ratio maximizing strategy (MS Cross) attains a mean return of 0.77% and a much lower volatility (1.92%), yielding a Sharpe ratio of 1.38. Consequently, a mean–variance investor would find the Sharpe-ratio-maximizing strategy considerably more attractive, whereas an investor

solely targeting expected returns would prefer the expected-return-maximizing strategy. Thus far, the out-of-sample performance aligns closely with the ex ante investment objectives.

Next, we consider a strategy that maximizes the Sharpe ratio using self-prediction to isolate the incremental contribution of cross-predictive relations relative to self-predictive relations. The key distinction between these two strategies lies in the structure of the connection matrix Ψ. Under cross-prediction, Ψ is a full 9×9 matrix, capturing all pairwise interactions among the characteristics and returns of assets. In contrast, under self-prediction, Ψ is restricted to its diagonal terms.

The second and third rows of Table 1 report the performance of the Sharpe-ratio–maximizing strategies under cross-prediction and self-prediction, respectively. The cross-prediction strategy (MS Cross) delivers a Sharpe ratio of 1.38 with a mean return of 0.77%, whereas the self-prediction variant (MS Self) achieves a lower Sharpe ratio of 1.02 and the lowest mean return of 0.43% . This gap in both risk-adjusted and absolute returns illustrates the incremental benefit of incorporating cross-predictive relationships beyond self-prediction alone, underscoring the pivotal role of cross-predictive dynamics in enhancing portfolio performance.

To provide further economic perspective on the value of accounting for cross-stock predictability, we compute the certainty equivalent return of the investment strategies. The certainty equivalent is defined as $CE = \mu - \frac{\gamma}{2}\sigma^2$, where μ and σ are the expected return and volatility of the strategy, respectively, and the risk aversion parameter γ is set to 2. Accounting for cross-predictability, the certainty equivalent rate of return is approximately 8.80% per year—3.76% higher than that of self-predictability—indicating economically significant gains.

We next maximize expected return and Sharpe ratio under the zero-cost and leverage-two constraints. The fourth and fifth rows of Table 1 report these constrained strategies, confirming that imposing the zero-cost restriction reduces expected returns for both objectives. Nevertheless, even with zero cost and fixed leverage, the Sharpe-ratio–maximizing strategy outperforms the expected-return–maximizing strategy, delivering a higher mean return (0.54% vs. 0.49%) and a substantially higher Sharpe ratio (1.26 vs. 0.53).

To provide additional insight into cross-prediction and self-prediction strategies, Table 2 reports the estimated values of Λ, Ψ, and ω for each approach without imposing the zero-cost constraint. The estimation window spans 120 months, from December 2003 to November 2023, cover-

ing our final out-of-sample period. Panel A presents the Sharpe-ratio-maximizing cross-prediction strategy; Panel B presents the Sharpe-ratio-maximizing self-prediction strategy; and Panel C reports the differences in the portfolio weights ω between the two.

In Table 2, Panel A shows that the estimated Λ coefficient for book-to-market equity (BM) is -0.08, whereas the coefficients for the other four characteristics are all positive, with the smallest value at 0.23. This suggests that the Sharpe-ratio–maximizing strategy with cross-prediction is well balanced across the five characteristics. The full 9×9 matrix Ψ exhibits substantial values both on and off the diagonal: the average absolute value of its diagonal entries is 0.036, compared to an average absolute off-diagonal entry of 0.081, indicating that cross-predictive relationships play an even more substantial role in defining the trading strategy.

Panel B of Table 2 shows that under self-prediction the estimated Λ coefficients exhibit greater dispersion—asset growth (INV) even turns negative—while Ψ is constrained to its diagonal (average absolute value of 0.25, all off-diagonals zero). This contrast underscores the structural impact of omitting cross-predictive terms.

Panel C reports how the optimal weights ω shift between cross- and self-prediction: under cross-prediction, long exposures to ME1 BM3 and ME3 BM1 increase, and shorts in ME1 BM1, ME2 BM1, and ME3 BM2 deepen. For example, the ME2 BM1 position is -0.26 under cross-prediction—driven by off-diagonal Ψ entries of -0.14 and -0.15—whereas it is substantially smaller under self-prediction.

As noted earlier, the optimal trading strategy that accounts for cross-predictability delivers a 3.76% higher certainty equivalent return, suggesting that the estimated Λ and Ψ, which determine the portfolio weights ω, differ to an economically significant degree when cross-predictability is incorporated, relative to the benchmark case of self-predictability.

In summary, the results in Tables 1 and 2 confirm that incorporating cross-predictive relationships is valuable for constructing robust investment strategies, even in a low-dimensional illustrative setting.

Table 1: Performance of Strategies of a Toy Example

This table reports the monthly average return, monthly standard deviation, and annualized Sharpe ratio for five strategies in a low-dimensional toy example involving five characteristics and nine assets. The strategies are:

1. Unconstrained expected-return maximization with cross-prediction;
2. Unconstrained Sharpe-ratio maximization with cross-prediction;
3. Unconstrained Sharpe-ratio maximization with self-prediction;
4. Zero-cost, leverage-two expected-return maximization with cross-prediction;
5. Zero-cost, leverage-two Sharpe-ratio maximization with cross-prediction.

	Mean	Std	Sharpe Ratio	Cost
MR Cross	1.96	21.89	0.31	Not Zero Cost
MS Cross	0.77	1.92	1.38	Not Zero Cost
MS Self	0.43	1.45	1.02	Not Zero Cost
MR Cross ZC	0.49	3.22	0.53	Zero Cost
MS Cross ZC	0.54	1.47	1.26	Zero Cost

Table 2: Estimates for Λ and Ψ of a Toy Example

This tables reports the estimated values for λ, Ψ, and ω of the maximizing Sharpe ratio strategies with cross-prediction in Panel A and self-prediction in Panel B. The Λ vector has five elements corresponding to five characteristics: ME, BM, OP, INV, and MOM. There are nine assets for investment: the three-by-three sorted portfolio on ME and BM. Specifically, they are ME1 BM1, ME1 BM2, ME1 BM3, ME2 BM1, ME2 BM2, ME2 BM3, ME3 BM1, ME3 BM2, ME3 BM3. The Ψ is a nine-by-nine matrix, where the element i,j corresponds to the strength of the predictive relationship of the asset i's signals to asset j's returns. For cross-prediction in Panel A, the Ψ has 81 values to estimate, while for self-prediction in Panel B, the Ψ is only active in 9 values in the diagonal. In addition, the following two rows of panels A and B report the absolute average of the diagonal and off-diagonal terms of Ψ. Finally, Panel C shows the change of ω from cross- to self-prediction strategies.

Panel A: Cross-Prediction

Λ	ME	BM	OP	INV	MOM				
	0.23	-0.08	0.34	0.65	0.64				
Ψ	0.01	0.16	-0.20	0.02	-0.05	-0.22	-0.16	0.00	-0.14
	-0.02	-0.03	-0.02	0.08	0.07	0.11	0.04	0.16	0.15
	0.32	0.17	-0.01	0.08	0.02	0.16	-0.23	0.03	0.03
	-0.17	0.10	0.03	0.14	0.06	-0.19	0.30	-0.05	-0.14
	-0.08	-0.08	0.10	0.01	0.00	0.00	0.14	0.04	0.03
	-0.10	-0.16	-0.08	-0.14	-0.10	0.02	-0.14	0.03	0.07
	0.22	-0.01	0.09	-0.15	-0.01	0.04	0.07	-0.19	-0.04
	-0.02	-0.04	0.09	-0.05	0.05	0.09	0.05	-0.02	0.08
	-0.15	-0.12	-0.01	0.01	-0.04	-0.01	-0.08	-0.01	-0.03
	Absolute Average of Diagonal Terms Ψ								0.036
	Absolute Average of Off-Diagonal Terms Ψ								0.081
ω	-0.17	-0.26	0.31	-0.24	0.01	-0.03	0.51	-0.30	-0.05

Panel B: Self-Prediction

Λ	ME	BM	OP	INV	MOM				
	0.37	-0.27	0.48	-0.17	0.73				
Ψ	0.21	0	0	0	0	0	0	0	0
	0	0.71	0	0	0	0	0	0	0
	0	0	-0.16	0	0	0	0	0	0
	0	0	0	0.18	0	0	0	0	0
	0	0	0	0	0.32	0	0	0	0
	0	0	0	0	0	-0.04	0	0	0
	0	0	0	0	0	0	0.53	0	0
	0	0	0	0	0	0	0	-0.12	0
	0	0	0	0	0	0	0	0	0.01
	Absolute Average of Diagonal Terms Ψ								0.25
	Absolute Average of Off-Diagonal Terms Ψ								0
ω	-0.04	-0.20	0.13	0.02	0.01	0.02	0.44	-0.07	0.00

Panel C: Change of Weights from Cross- to Self-Prediction

ID	ME1 BM1	ME1 BM2	ME1 BM3	ME2 BM1	ME2 BM2	ME2 BM3	ME3 BM1	ME3 BM2	ME3 BM3
$\Delta\omega$	-0.13	-0.06	0.18	-0.26	0.00	-0.05	0.06	-0.23	-0.05

Table 3: Performance of Cross-Predictive Strategies

This table reports the monthly average return, monthly standard deviation, and annualized Sharpe ratio of cross-predictive strategies. The strategies are zero-cost and leverage two. MR and MS are strategies to maximize expected return and Sharpe ratio, respectively. Panels A and C are for investing in 138 spread portfolios, and Panels B and D are for 544 bi-variate sorted portfolios. In Panels A and B, we report the results of the whole out-of-sample period from February 1973 to December 2023 and the high and low sentiment periods split by the sentiment median value over the sample periods from February 1973 to December 2023. In Panels C and D, we report for January 1990 to December 2023, and the high and low VIX periods split by the VIX median value over the sample periods from 1990 to 2023.

	1973:02-2023:12			SENT High			SENT Low		
	μ	σ	SR	μ	σ	SR	μ	σ	SR
Panel A: Spread Portfolios									
MR	0.42	3.23	0.45	0.73	3.79	0.67	0.11	2.53	0.15
MS	0.18	0.28	2.19	0.17	0.28	2.15	0.18	0.28	2.22
Panel B: BiSort Portfolios									
MR	0.45	3.02	0.52	0.48	3.35	0.49	0.42	2.66	0.54
MS	0.18	0.17	3.62	0.21	0.18	3.96	0.16	0.17	3.32
	1990:01-2023:12			VIX High			VIX Low		
	μ	σ	SR	μ	σ	SR	μ	σ	SR
Panel C: Spread Portfolios									
MR	0.33	3.83	0.30	0.59	4.97	0.41	0.07	2.14	0.12
MS	0.15	0.29	1.79	0.16	0.35	1.63	0.13	0.21	2.19
Panel D: BiSort Portfolios									
MR	0.39	3.20	0.42	0.57	3.87	0.51	0.20	2.33	0.30
MS	0.17	0.19	3.17	0.20	0.21	3.25	0.15	0.16	3.23

5.2 Zero-Cost Linear Strategies

Table 3 reports the performance of linear cross-predictive strategies implemented as zero-cost, leverage-two portfolios, comparable to common factor and anomaly implementations. MR and MS denote the strategies that maximize expected return and the Sharpe ratio, respectively. In Panel A, we consider an investment universe with 138 spread portfolios detailed in the data section. Over the full sample period, MR achieves a monthly average return of 0.42% with an annualized Sharpe ratio of 0.45, whereas MS records a lower monthly average return of 0.18% but a substantially higher annualized Sharpe ratio of 2.19.

We further analyze performance during evolving market states by splitting the out-of-sample period at the median of the investor sentiment index (Baker and Wurgler, 2006).[3] During high-sentiment regimes, MR delivers an average monthly return of 0.73%, while in low-sentiment regimes its return falls to 0.11%. The MS strategy exhibits robust Sharpe ratios across both regimes: 2.15 in high-sentiment periods and 2.22 in low-sentiment periods.

In Panel B, we evaluate investments in 544 bi-variate sorted portfolios as detailed in the data section. Over the full out-of-sample period (January 1973–December 2023), MR delivers a monthly average return of 0.45% and an annualized Sharpe ratio of 0.52, whereas MS achieves an exceptionally high annualized Sharpe ratio of 3.62. In sub-period analyses, MR's average return increases during high-sentiment regimes, while MS maintains Sharpe ratios above 3 in both high- and low-sentiment periods.

In Panels C and D, we split the period January 1990–December 2023 at the median of the VIX index.[4] In Panel C (spread portfolios), MR's return is 0.59% during high-VIX regimes and 0.07% during low-VIX regimes (0.33% full sample), while MS records Sharpe ratios of 1.63 and 2.19 in high- and low-VIX regimes (1.79 full sample).

In Panel D (bi-variate sorted portfolios), MR attains a monthly average return of 0.39% and an annualized Sharpe ratio of 0.42, while MS achieves a Sharpe ratio of 3.17. MR's return remains higher in high-VIX regimes, and MS sustains Sharpe ratios above 3.23 in both high- and low-VIX

[3]The sentiment data spans July 1965 to December 2023 and is obtained from the variable 'SENT" on Jeffrey Wurgler's website: https://pages.stern.nyu.edu/~jwurgler/data/SENTIMENT.xlsx.

[4]The VIX data spans 1990 to 2023 and is obtained from the CBOE: http://www.cboe.com/products/vix-index-volatility/vix-options-and-futures/vix-index/vix-historical-data/.

regimes.

In unreported analyses, we derive the maximum Sharpe ratio strategy under an economically motivated restriction: we constrain the Ψ matrix such that characteristics of large-cap stocks can predict returns of small-cap stocks, but not the reverse. While this restriction leads to only a marginal improvement in the out-of-sample Sharpe ratio, it is possible that alternative economically grounded constraints on Ψ could yield greater benefits. We leave a more comprehensive exploration of such structures to future research.

In summary, MR strategies deliver high expected returns during high-sentiment and high-VIX regimes, but considerably lower expected returns otherwise. In contrast, MS strategies consistently achieve superior Sharpe ratios across all market states.

5.2.1 Comparing with Principal Portfolios (PP)

We compare the principal-portfolio-based trading strategies of Kelly et al. (2023) with our own over the out-of-sample period from 1973 to 2019, as in their study. The results are reported in Table 4.

Panel A, row 1 (PP–ME), reports the performance of the first principal portfolio on the market-equity signal: a 3.27% monthly expected return, a 0.51 annualized Sharpe ratio, and a sum of absolute equity positions equal to 23.22. Rows 2 and 3 present the first principal portfolios for the book-to-market and momentum signals, which achieve Sharpe ratios of 0.60 and 0.48, respectively, with similarly high leverage. Averaging the first principal portfolios across all 138 signals yields the PPEW strategy, which delivers a 2.83% monthly expected return and a 0.56 annualized Sharpe ratio. Notably, the leverage of PPEW is only 1.35, reflecting substantial diversification benefits by equal weighted average across predictors.

Our maximum-expected-return strategy achieves an 11.55% monthly expected return and a 0.51 annualized Sharpe ratio, with leverage of 45.94. Overall, the maximum-expected-return strategy slightly underperforms the Principal Portfolios, albeit remains reasonably close to them.

By contrast, the MS strategy harnesses multiple predictors to diversify exposures and optimize risk-adjusted returns, achieving an annualized Sharpe ratio of 2.37 with a leverage factor of 33.46. While the PP approach targets expected return subject to a volatility constraint, our strategy is de-

rived directly from Sharpe ratio maximization. As a result, it places greater emphasis on balancing return and risk, leading to improved performance on risk-adjusted metrics in our empirical setting.

To enhance implementability, we impose zero net-cost and leverage-two constraints on both strategies. Panel B of Table 4 reports the resulting performance. Under these constraints, the maximum-expected-return strategy (Row 1) achieves a 0.46% monthly expected return and an annualized Sharpe ratio of 0.51, while the maximum-Sharpe-ratio strategy (Row 2) attains a 0.19% monthly expected return and an annualized Sharpe ratio of 2.36. In both cases, the portfolios maintain zero net cost and a constant leverage of two in every period.

Overall, the maximum-Sharpe-ratio strategy remains highly competitive—delivering superior risk-adjusted performance relative to a range of recent approaches, including Principal Portfolios. Accordingly, we focus our subsequent analyses to the constrained max-SR strategy.

Table 4: A First Comparison on the Performance of PP, MR, and MS

This table reports each strategy's monthly average return (%), monthly standard deviation (%), annualized Sharpe ratio, time-series average net exposure to basic assets, and time-series average gross exposure to basic assets. **PP-ME** is the pure-play (PP) strategy using the market-equity signal; **PP-BM** uses book-to-market; **PP-MOM** uses momentum. **PP-EW** is an equal-weighted combination of the three PP strategies. **MR** is our maximum-expected-return strategy, and **MS** is our maximum-Sharpe-ratio strategy. Panel A places no leverage or cost constraints on the strategies. Panel B imposes two constraints—a zero-cost requirement and a leverage restriction—on all strategies. Data span January 1963 through December 2019 (from PP's replication package on the *Journal of Finance* website), with the out-of-sample period running from February 1973 to December 2019.

Strategy	Avg%	Std%	SR	Sum	ASum
Panel A: Linear Strategies					
PP-ME	3.27	22.32	0.51	23.21	23.22
PP-BM	4.64	26.94	0.60	12.62	14.45
PP-MOM	3.65	26.41	0.48	23.94	25.34
PPEW	2.83	17.52	0.56	1.03	1.35
MR	11.55	76.52	0.52	8.19	45.94
MS	3.13	4.56	2.37	1.06	33.46
Panel B: Linear Strategies with Zero Cost					
MR	0.46	3.10	0.51	0.00	2.00
MS	0.19	0.27	2.36	0.00	2.00

5.2.2 Cross-Prediction SDF versus Self-Prediction SDF

The existing literature on SDF estimation has predominantly focused on self-predictive frameworks, where each asset's signals are used solely to forecast its own returns. Kozak et al. (2020) construct a robust SDF formed from a small number of principal components of characteristics-based factors. Kelly et al. (2019) propose Instrumented PCA with the belief that the factor loadings on SDF are determined by assets characteristics, overcoming the limitations of static loading in PCA. Lettau and Pelger (2020) find that the SDF estimated on Risk-Premium PCA is more highly correlated with the true SDF than those estimated on PCA. Luo et al. (2025) estimate the SDF with observable characteristics-based factors with L1-penalized SDF regression; whereas, Didisheim et al. (2024) apply the L2-penalized SDF regression on observable and Random-Fourier-Feature generated factors. All of these papers have been working on high-dimensional characteristics-based portfolios to estimate the SDF, where the belief of self-prediction are embedded the portfolios.

By contrast, our framework utilizes managed portfolios inherently reflecting the belief of cross-prediction: π_s (6), χ_Φ (19), and χ_Λ (22). Whether cross-predictive strategies—where signals from one asset help predict the returns of others—can systematically outperform self-predictive ones in high-dimensional settings remains an open question.

To investigate this, we construct the self-predictive strategy by restricting the matrix Ψ to its diagonal, thereby eliminating all cross-asset interactions. Conceptually, this restriction renders self-prediction a special case of the economic constraints introduced in Section 2.3, in which each asset's signals are permitted to forecast only its own returns.

Panel A of Table 5 reports the empirical performance of the Sharpe-ratio–maximizing strategies on the 138 spread portfolios. The self-predictive strategy achieves a Sharpe ratio of 1.80, while the cross-predictive counterparts attain Sharpe ratios of 2.19, both with and without zero-cost and leverage-two constraints. This 0.39 differential underscores the incremental value of incorporating cross-asset predictive signals.

Panel B reports results for the 544 bi-variate sorted portfolios. The self-predictive strategy achieves a Sharpe ratio of 2.51, while the cross-predictive variants reach 3.62 and 3.55 under constrained and unconstrained implementations, respectively. This gap of more than 1.00 in Sharpe ratio highlights the significant contribution of off-diagonal elements in Ψ to improved portfolio

performance.

Overall, the evidence confirms that cross-predictive strategies materially enhance the estimation and performance of stochastic discount factors—particularly in richer portfolio universes and longer samples.

Table 5: Cross- vs Self-Prediction

This table reports the monthly average returns (%), monthly standard deviation(%), and annualized Sharpe ratio, time-series average of the sum of positions on basic assets, and time-series average of the absolute sum of positions on basic assets. The objective of the strategies is to maximize the Sharpe ratio. Notably, the cross-prediction strategies can be solved with and without the zero-cost constraint; however, the self-prediction strategy does not have an analytic solution once adding the zero-cost constraint, see Section 2.3 for details. Panel A invests on spread portfolios and Panel B is for bi-variate sorted portfolios. The out-of-sample period is February 1973 to December 2023.

	μ	σ	SR	Sum	ASum
Panel A: Spread Portfolios					
MS Self	0.72	1.38	1.80	0.64	3.80
MS Cross	2.98	4.70	2.19	1.04	33.70
MS Cross ZC	0.18	0.28	2.19	0.00	2.00
Panel B: BiSort Portfolios					
MS Self	1.09	1.50	2.51	0.18	4.04
MS Cross	10.95	10.67	3.55	1.40	90.23
MS Cross ZC	0.18	0.17	3.62	0.00	2.00

5.2.3 Factor Spanning Tests

We conduct a series of factor-spanning tests to assess whether established asset pricing factors fully explain the expected returns of the maximum-Sharpe-ratio (MS) strategy. Table 6 reports monthly alphas, factor loadings, and associated t-statistics. Panel A (MS Spread) presents results for the spread portfolios, while Panel B (MS BiSort) reports findings for the bivariate sorted portfolios.

We first evaluate the Fama and French (2015) five-factor model (FF5). The MS Spread portfolio exhibits a modest loading on SMB ($\beta = 0.01$, $t = 2.07$) but insignificant exposures to the other four factors, while delivering a highly significant monthly alpha of 0.18% ($t = 15.13$). This sug-

gests that the strategy's returns are largely orthogonal to the FF5 factors. Also, we augment the FF5 model with short-term reversal (REV), momentum (UMD), and liquidity (LIQ) factors (Pastor and Stambaugh, 2003). In this expanded specification, MS Spread shows significant loadings on REV and UMD but not on LIQ, while its alpha remains economically and statistically significant at 0.16% ($t = 13.96$). These findings indicate that reversal and momentum effects partially explain the strategy's performance, with little role for liquidity risk.

Next, we then examine the Hou et al. (2015) four-factor model, which incorporates investment (R_IA) and profitability (R_ROE) factors alongside market and size. MS Spread displays negligible loadings on R_IA and R_ROE, while maintaining a highly significant alpha. The Stambaugh and Yuan (2017) mispricing factors—managerial ownership (MGMT) and performance (PERF)—also fail to subsume the strategy's returns: MS Spread shows minimal exposures to both factors, with an alpha of 0.17% ($t = 13.06$). Then, we assess the Daniel et al. (2020) model, which includes post-earnings announcement drift (PEAD) and financing (FIN) factors. While MS Spread loads significantly on PEAD, its alpha remains robust at 0.17% ($t = 13.42$), and it shows no meaningful exposure to FIN. Finally, in a comprehensive regression incorporating all fourteen factors, MS Spread maintains an alpha of 0.16% ($t = 11.62$), with statistically significant but economically small loadings on SMB, UMD, REV, LIQ, FIN, and R_IA. These results collectively demonstrate that the MS strategy's expected returns cannot be fully explained by existing factor models.

Panel B corroborates these findings. The MS BiSort portfolio displays statistically significant but economically modest loadings on RMW, CMA, REV, and PERF. Notably, it maintains a monthly alpha of 0.17% ($t = 20.45$) even after controlling for all fourteen factors, further supporting the strategy's robustness to established factor models.

In summary, across all specifications—including the Fama–French five-factor model with REV, UMD, and LIQ augmentations, as well as the Hou–Xue–Zhang, Stambaugh–Yuan, and Daniel–Hirshleifer–Sun frameworks, and even the comprehensive fourteen-factor regression—the MS Spread and MS BiSort portfolios exhibit persistently large and highly significant alphas with only moderate loadings on existing factors. This suggests that conventional models may miss the cross-stock return predictability captured by our strategy. Below, we further analyze the pricing content of the maximum Sharpe ratio portfolio.

Table 6: Alpha and Factor Loadings

This table reports the alphas, factor loadings, and their t-values (in parentheses) obtained from the factor-spanning tests of regressing the strategy returns on other asset pricing factors. We have scaled the original strategy and factor returns by 100 for percentage compatibility, aiding coefficient comparability. This table focus on the maximizing Sharpe ratio strategy with zero-cost and leverage two. Panel A (MS Spread) displays the results for investing in spread portfolios, while Panel B (MS BiSort) shows for bi-variate sorted portfolios. The factors include FF5 factor, momentum factor (UMD), short-term reversal factor (REV), liquidity factor (LIQ) from Pastor and Stambaugh (2003), short-horizon inattention factor (PEAD) and long-horizon financing factor (FIN) from Daniel et al. (2020), investment factor (R_IA) and return on equity factor (R_ROE) from Hou et al. (2015), management factor (MGMT) and performance factor (PERF) from Stambaugh and Yuan (2017). PEAD and FIN are available before December 2018. MGMT and PERF are available before December 2016. All other factors are available during the whole sample period. One, two, and three asterisks indicate significance at the 10%, 5%, and 1% levels, respectively.

Alpha	Market	SMB	HML	RMW	CMA	UMD	REV	LIQ	PEAD	FIN	R_IA	R_ROE	MGMT	PERF
						Panel A: MS Spread								
0.18***	-0.00	0.01**	-0.00	-0.00	0.01									
(15.13)	(-1.47)	(2.07)	(-0.59)	(-0.89)	(0.68)									
0.16***	-0.00	0.01*	0.00	-0.01	0.00	0.02***	0.01***	-0.00						
(13.96)	(-0.99)	(1.85)	(0.57)	(-1.18)	(0.20)	(5.77)	(3.36)	(-1.36)						
0.17***	-0.00								0.02***	0.00				
(13.42)	(-0.00)								(3.14)	(0.78)				
0.18***	-0.00*	0.01***									-0.00	0.00		
(14.72)	(-1.65)	(2.60)									(-0.51)	(0.59)		
0.17***	0.00	0.01***											0.01*	0.01***
(13.06)	(0.78)	(3.18)											(1.86)	(4.05)
0.16***	0.00	0.01*	-0.00	-0.01	0.02	0.01**	0.01***	-0.00**	0.01	0.01**	-0.03*	-0.01	0.01	0.01
(11.62)	(1.09)	(1.77)	(-0.56)	(-0.73)	(1.17)	(2.48)	(3.62)	(-2.09)	(1.19)	(1.98)	(-1.82)	(-1.62)	(0.58)	(1.60)
						Panel B: MS BiSort								
0.18***	-0.00	0.00	0.00	0.02***	-0.00									
(24.44)	(-0.42)	(1.08)	(1.19)	(4.96)	(-0.13)									
0.17***	-0.00	0.00	0.01**	0.02***	-0.00	0.01***	0.01***	-0.00						
(23.33)	(-0.17)	(0.78)	(2.11)	(4.87)	(-0.47)	(5.40)	(4.01)	(-1.50)						
0.18***	0.00								0.01**	0.01***				
(22.70)	(1.45)								(2.02)	(4.86)				
0.17***	-0.00	0.00									0.01*	0.01***		
(23.40)	(-0.20)	(1.04)									(1.75)	(4.26)		
0.18***	0.00**	0.00											0.01***	0.01***
(23.12)	(2.12)	(0.78)											(3.90)	(5.43)
0.17***	0.00	0.00	0.00	0.02***	-0.02**	0.00	0.01***	-0.00	0.01*	0.00	0.02*	-0.01	0.01*	0.01**
(20.45)	(1.35)	(1.27)	(0.41)	(2.72)	(-2.07)	(1.32)	(5.22)	(-1.04)	(1.92)	(0.23)	(1.77)	(-1.64)	(1.90)	(2.44)

5.2.4 Evolution of Sharpe Ratios over the Sample Period

To assess the persistence and evolution of risk-adjusted returns over time, Figure 1 Panel A shows the five-year trailing Sharpe ratios of our two MS strategies alongside those of the market and momentum factors. By smoothing over a half-decade window, we can observe how the trading strategies respond to changing market conditions.

Both MS strategies deliver eye-catching Sharpe ratios in the 1990s—MS BiSort climbs as high as 4–7 before 2000, and MS Spread approaches 3—reflecting their ability to capture persistent value-enhancing opportunities. After 2000, however, it is natural to see some attenuation: wider adoption of anomaly tradings, increased market liquidity, and a lower-volatility regime tend to compress excess returns over time. Accordingly, by the end of 2023, MS BiSort's trailing Sharpe has moderated to about 1.5 and MS Spread to roughly 1.

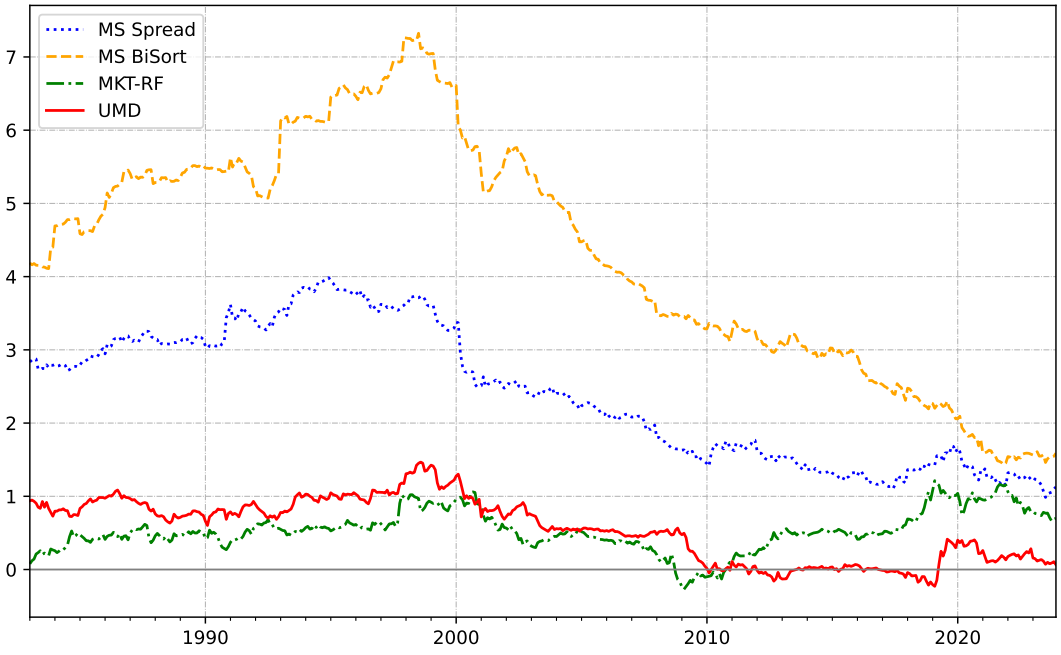
To make more clear comparison, Figure 1 Panel B shows the Sharpe ratio of each strategy relative to that of MS BiSort. In early sample before 2000, the Sharpe ratio of MS Spread is approximately 60% that of MS BiSort, and market and momentum factors have below 20% Shape ratio relative to MS BiSort. In the most recent sample, the Sharpe ratios of MS Spread, MKT-RF, and UMD are 71%, 44%, and 4% that of MS BiSort.

For context, both the market factor's rolling Sharpe ratio and that of the momentum factor remain well below our strategies over the entire forty-year span. Although the performance gap narrows in the post-2000 era, both MS strategies continue to deliver robust risk-adjusted returns relative to these broad benchmarks.

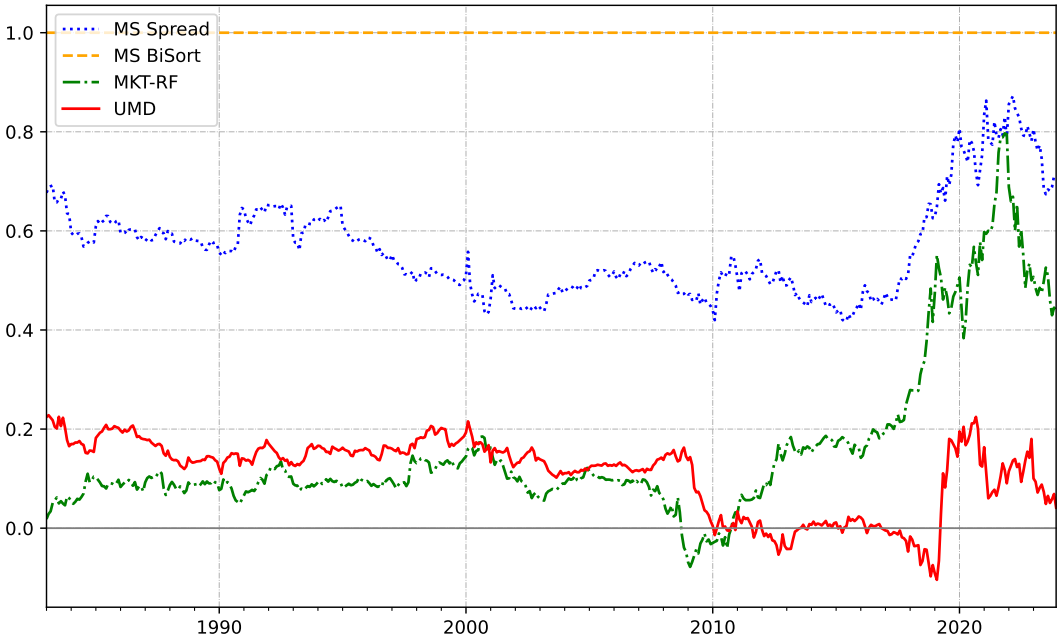
Table 7 reports the (annualized) Sharpe ratios of the cross-predictive investment strategies, Fama-French five factors, and momentum factor for three sample periods: the whole OOS period from 1973:02 to 2023:12, before 2000:01, and after 2000:01. Although, the Sharpe ratios of our strategies attenuate after 2000, they remain competitive compared to the benchmark factors in three sample periods.

Figure 1: Sharpe Ratio of Strategies

The figure depicts five-year trailing (annualized) Sharpe ratios for the cross-predictive investment strategies. Panel A shows the Sharpe ratio, while Panel B shows the Sharpe ratio of each strategy divided by that of "MS BiSort." "MS Spread" is the strategy to maximize Sharpe ratio investing in the spread portfolios."MS BiSort" is the strategy to maximize Sharpe ratio investing in the bi-variate sorted portfolios. The out-of-sample period is from February 1973 to December 2023 in monthly frequency, and the first five-year Sharpe ratio is obtained for January 1983. For comparison, the market factor (MKT-RF) and momentum factor (UMD) are included.



(a) Sharpe Ratio



(b) Sharpe Ratio Relative to MS BiSort

Table 7: Comparing Sharpe Ratios with Prevailing Factors

The table reports the (annualized) Sharpe ratios of the cross-predictive investment strategies, Fama-French five factors, and momentum factor. Three sample periods are 1973:02 to 2023:12, 1973:02 to 1999:12, and 2000:01 to 2023:12.

	MS Spread	MS BiSort	MKT-RF	SMB	HML	RMW	CMA	UMD
1973-2023	2.19	3.62	0.45	0.18	0.33	0.45	0.50	0.45
1973-1999	3.23	5.46	0.48	0.15	0.48	0.36	0.58	0.96
2000-2023	1.33	2.42	0.41	0.20	0.20	0.54	0.43	0.09

5.3 Signal Importance

To understand the economic underpinnings of our Sharpe-maximizing strategy, we examine the estimated values of Λ, which assign weights to firm-level predictive signals. These weights reflect the relative contribution of each signal to the stochastic discount factor (SDF). We focus on the absolute value of these weights averaged over time to assess long-term signal importance. Table 8 presents the ten most influential signals, ranked by their time-series average of absolute $|\Lambda|$ values, where Panel A is for spread portfolios and Panel B is for bi-variate sorted portfolios.

Panel A indicates that the most important signals are concentrated in the *investment* and *value* categories, investing in spread portfolios. For instance, the top signal—*liquidity of book assets* (Ortiz-Molina and Phillips, 2014)—receives an average importance of 0.141, while *dividend yield* (Litzenberger and Ramaswamy, 1979), the leading signal in the value theme, ranks seventh overall with an importance of 0.124. These findings suggest that the strategy places greater emphasis on firm fundamentals linked to capital structure, financing constraints, and valuation, rather than technical or return-based indicators.

As for Panel B, *profitability* dominates the top ten signals, followed by *size*, *low leverage*, and *low risk* themes. For instance, *return on equity* (Haugen and Baker, 1996), *operating profitability-to-lagged book equity* (Fama and French, 2015), and *profit margin* (Soliman, 2008) are top three signals, all belonging to *profitability*. Besides, *price per share* (Miller and Scholes, 1982) and *Amihud illiquidity* (Amihud, 2002) emerge from the *size* theme, recalling stronger size effects in the test assets sorted on size and other signals.

Table 8: Top Ten Signals by Λ

This table reports the top ten most important signals, Panel A for spread portfolios and Panel B for bi-variate sorted portfolios. The columns are abbreviation, theme, time-series average of absolute Λ, full name, and original publication of the signals. There are 13 themes following Jensen et al. (2023).

	Abbreviation	Theme	Λ	Full Name	Publication
			Panel A: Spread Portfolio		
2	aliq_at	Investment	0.141	Liquidity of book assets	Ortiz-Molina and Phillips (2014)
44	emp_gr1	Investment	0.129	Hiring rate	Belo et al. (2014)
116	sale_gr3	Investment	0.128	Sales growth (3 years)	Lakonishok et al. (1994)
9	be_gr1a	Investment	0.127	Change in common equity	Richardson et al. (2005)
24	col_gr1a	Investment	0.127	Change in current ope. lia.	Richardson et al. (2005)
115	sale_gr1	Investment	0.126	Sales growth (1 year)	Lakonishok et al. (1994)
34	div12m_me	Value	0.124	Dividend yield	Litzenberger and Ramaswamy (1979)
7	at_me	Value	0.122	Assets-to-market	Eugene and French (1992)
45	eq_dur	Value	0.121	Equity duration	Dechow et al. (2004)
6	at_gr1	Investment	0.121	Asset growth	Cooper et al. (2008)
			Panel B: BiSort Portfolio		
71	ni_be	Profitability	0.142	Return on equity	Haugen and Baker (1996)
86	ope_bel1	Profitability	0.137	Ope. profits-to-lagged be	Fama and French (2015)
42	ebit_sale	Profitability	0.137	Profit margin	Soliman (2008)
90	prc	Size	0.136	Price per share	Miller and Scholes (1982)
77	o_score	Profitability	0.136	Ohlson O-score	Dichev (1998)
85	ope_be	Profitability	0.135	Operating profits-to-be	Fama and French (2015)
16	bidaskhl_21d	Low Leverage	0.132	The high-low bid-ask spread	Corwin and Schultz (2012)
41	ebit_bev	Profitability	0.131	Return on net operating assets	Soliman (2008)
58	ivol_capm_252d	Low Risk	0.127	Idio. vol. to CAPM (21 days)	Ali et al. (2003)
4	ami_126d	Size	0.127	Amihud Measure	Amihud (2002)

Figure 2 presents the importance measures for all 138 signals, organized into 13 thematic categories (as defined in the Data section). Sub-figures (a) and (b) display theme-level importance for spread portfolios and bivariate-sorted portfolios, respectively.[5] The heatmap visualization employs color intensity to indicate importance levels—with red (blue) representing high (low) importance–allowing clear identification of which signals consistently influence portfolio construction.

In sub-figure A for spread portfolios, *investment*- and *value*-related signals dominate the red spectrum, reinforcing the role of tangible firm fundamentals. In contrast, *momentum*, *profit growth*, *debt issuance*, *seasonality*, and *short-term reversal* appear consistently in the blue range, indicating minimal weight in the Sharpe-maximizing SDF.

Turning to sub-figure B for bivariate-sorted portfolios, the *profitability* theme dominates the heatmap, particularly following a pronounced regime shift in the late 1980s. The *size* theme exhibits persistent importance throughout the sample period, reflecting the strong cross-sectional dispersion in firm size within our test assets. In contrast, *accruals*, *profit growth*, *seasonality*, and *short-term reversal* show consistently negligible predictive power over the entire sample.

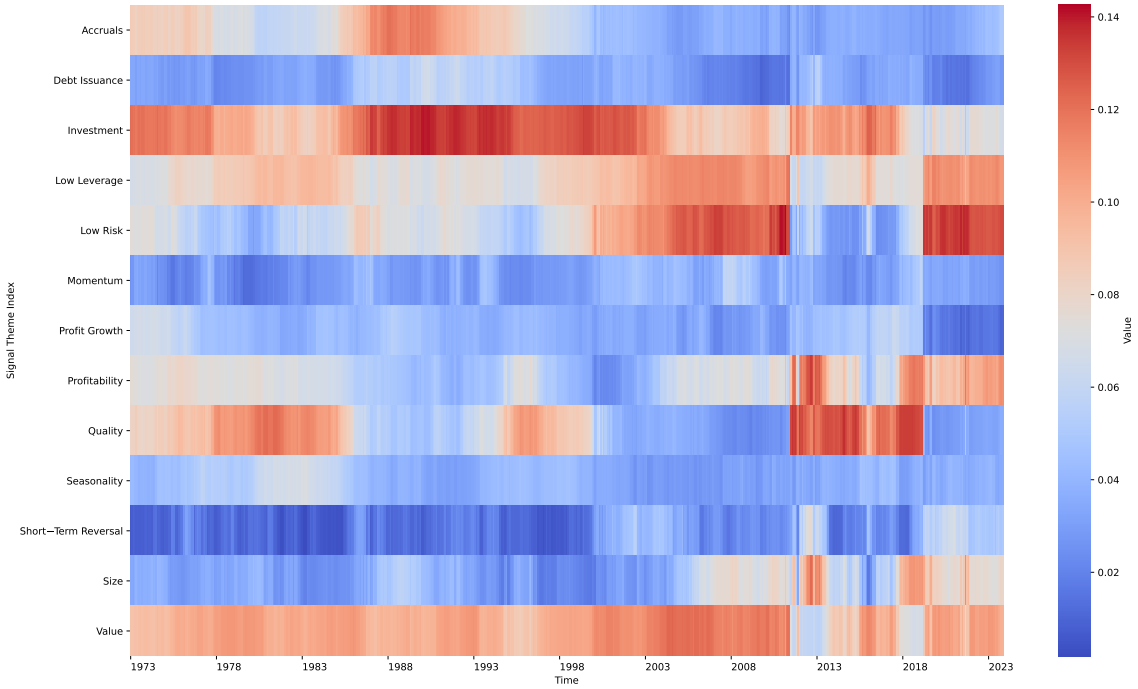
Our analysis reveals that while the dominant predictive role of *investment*, *value*, *profitability*, and *size* themes remains stable over time, certain signals—particularly *accruals* and *quality*—exhibit heightened importance during high-volatility or low-sentiment periods. This time variation suggests dynamic shifts in return predictability patterns, which our framework successfully captures through its adaptive structure.

In summary, our signal importance analysis demonstrates that the cross-predictive SDF is primarily driven by *stable, economically grounded predictors*, with negligible dependence on transient or noisy effects. These findings not only underscore the robustness and economic interpretability of our framework but also open new avenues for investigating the fundamental drivers of cross-sectional return predictability.

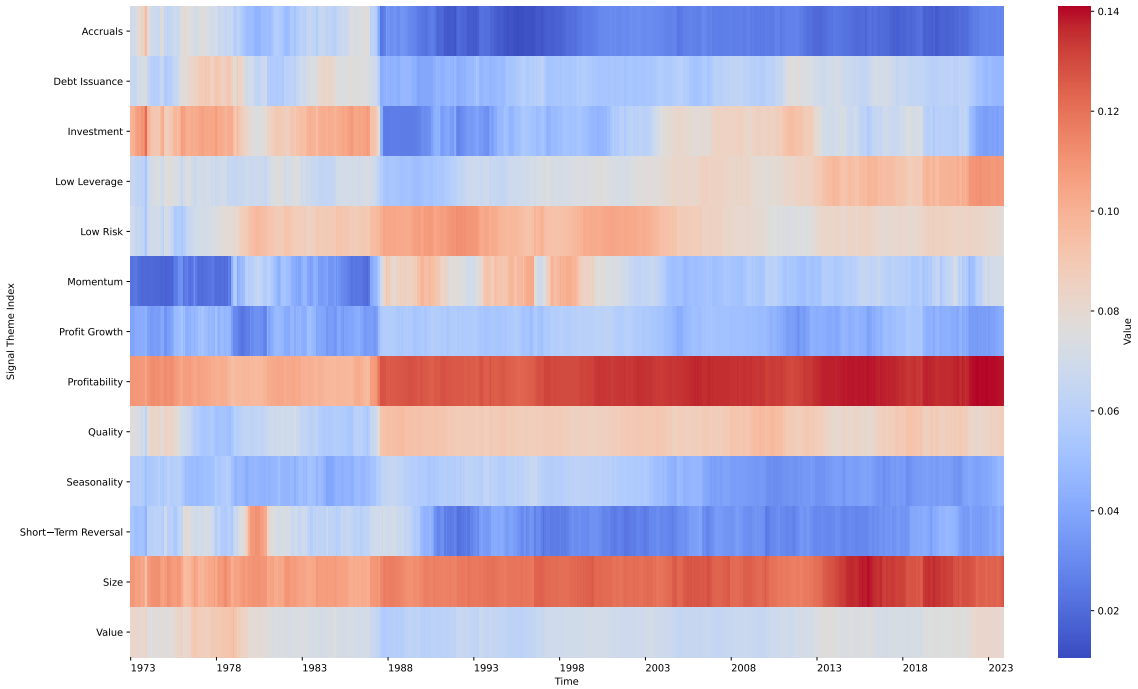
[5] We provide the time-varying signal-level importance measures in Figure E.1 of the Appendix.

Figure 2: Signal Importance

This figure depicts the heatmaps of signal importance Λ for each rolling-window estimation. Sub-figures A and B are for spread portfolios and bi-variate sorted portfolios, respectively. For interpretation, we focus on the absolute value of elements in Λ. We calculate the theme-level importance as the average of all signal-level importance within each theme. There are 13 themes following Jensen et al. (2023).



(a) Spread Portfolio: Theme-Level



(b) BiSort Portfolio: Theme-Level

5.4 Networks in the Cross Section

To uncover the economic structure embedded in the cross-predictive matrix Ψ, we interpret Ψ as the adjacency matrix of a directed network across N assets. This representation enables us to move beyond aggregate portfolio-level effects and examine how predictive information flows through the cross-section—identifying assets that function as net transmitters or receivers of signals and assessing the alignment of these linkages with economic groupings such as firm size.

Following the connectedness methodology of Diebold and Yılmaz (2014), we compute four summary statistics for each asset: *outgoing connectedness* (FROM), *incoming connectedness* (TO), *net connectedness* (NET), and *overall network intensity* (TOTAL). These measures help pinpoint dominant sources of predictability and inform the imposition of economically motivated sparsity structures on Ψ to improve interpretability and out-of-sample performance.

Let $\Psi_{i,j}$ denote the predictive influence of asset i on asset j. We define the network metrics as follows:

$$\text{FROM}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} |\Psi_{i,j}|, \tag{24}$$

$$\text{TO}_j = \sum_{\substack{i=1 \\ i \neq j}}^{N} |\Psi_{i,j}|, \tag{25}$$

$$\text{NET}_k = \text{TO}_k - \text{FROM}_k, \tag{26}$$

$$\text{TOTAL} = \frac{1}{N} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} |\Psi_{i,j}|. \tag{27}$$

Here, FROM_i measures the total strength of predictive signals sent from asset i to others, capturing how much i contributes to forecasting the returns of other assets. TO_j measures the total strength of predictive signals received by asset j from all other assets, reflecting how much j is influenced by the rest of the network. NET_k is the difference between incoming and outgoing connectedness, indicating whether asset k is a net transmitter (< 0) or net receiver (> 0) of predictive information. TOTAL aggregates the overall off-diagonal magnitude of Ψ across all asset pairs, summarizing the average intensity of cross-asset predictive linkages in the network. The use

of absolute values follows Diebold and Yılmaz (2014) and ensures all measures are non-negative, thereby capturing signal strength regardless of sign.

We compute these metrics monthly for two asset universes—138 spread portfolios and 544 bi-variate sorted portfolios—over $T = 611$ months. To investigate the firm-level characteristics driving variation in connectedness, we estimate monthly cross-sectional regressions:

$$\text{Connectedness}_{i,t} = \alpha_t + \boldsymbol{\beta}' \text{Char}_{i,t} + \varepsilon_{i,t}, \tag{28}$$

where $\text{Connectedness}_{i,t}$ is one of FROM_i, TO_i, or NET_i, and $\text{Char}_{i,t}$ is a vector of observable characteristics. We report time-series averages of the estimated coefficients along with Newey–West (Newey and West, 1987) t-statistics using a Bartlett kernel and lag length $L = 4(T/100)^{2/9} \approx 6$.

Table 9 reports the results of monthly cross-sectional regressions of three network connectedness measures—FROM, TO, and NET—on firm characteristics for two groups of test assets: spread portfolios (Panel A) and bi-variate sorted portfolios (Panel B). The results reveal economically intuitive patterns linking a stock's network role to size, valuation, profitability, investment, momentum, and several trading frictions.

In Panel A for spread portfolios, the FROM regressions, measuring how much a stock helps predict others, we observe that smaller stocks (low ME), high book-to-market (BM), high profitability (OP), and high momentum (MOM) stocks tend to transmit stronger signals to others. These firms—small, value, profitable, and past winners—have greater forecasting influence, possibly because they aggregate market-wide information or drive co-movements. Additionally, stocks with low illiquidity (ILL) and low turnover (TRN) exhibit higher FROM, suggesting that liquidity increase a stock's impact to the network. Volatility (VLT), by contrast, enters positively, implying that more volatile stocks spill predictive attention. Notably, the coefficient on size (ME) becomes insignificant, once controlling five trading frictions, which means that the size effect on FROM is a manifestation of trading frictions but not size itself.

The TO regressions, which capture how strongly a stock is predicted by others, show the opposite patterns on many characteristics. Stocks with high ME, high BM, low OP, low INV, high MOM, high ILL, high VLT, and high BETA receive more predictive inputs from others. This suggests that firms that are large, volatile, illiquid, and priced as value stocks appear more susceptible

to being forecasted using cross-sectional information. Interestingly, high-MOM stocks both receive and transmit signals, indicating they may act as informational amplifiers within the network.

The NET regressions, defined as TO − FROM, consolidate these effects to identify whether a stock is a net receiver or transmitter of predictive information. Stocks that are large (high ME), low-BM, low-OP, low-INV, low-MOM, high-ILL, and high-TRN tend to be net receivers, while small, value, profitable, non-investing, and momentum-driven stocks are net transmitters. These directional patterns highlight a persistent asymmetry: small, liquid, value, strong profitability, and conservative investment firms disseminate predictive signals, while larger and illiquid firms absorb them. The NET regressions also reveal that turnover (TRN) consistently distinguishes transmitters from receivers, suggesting that actively traded stocks play a key role in receiving predictive information.

In Panel B for bi-variate sorted portfolio, these patterns shift dramatically. For ease of interpretation, we focus on the NET regressions. We find that small, high-BM, high-OP, high-INV, and high-MOM firms are net receivers in the network, while big, growth, weak-profitable, active-investing, and low-momentum stocks are net transmitters. After controlling five trading frictions in the regressions, the coefficient on size become significantly positive, while other four coefficients are still positive. As for trading frictions, stocks with low volume, high turnover, and low market-beta tend to receive spillovers from others than transmitting signals to others.

Together, two sets of test assets demonstrate significant correlations between network connectedness and asset characteristics, shedding light on that the determinants of cross-asset spillover effects. The estimated Ψ matrix embeds an economically interpretable hierarchy of signal flows, shaped by firm fundamentals and market frictions. This structure supports imposing sparsity or blockwise restrictions to enhance interpretability and control overfitting—especially by limiting signal flows that contradict observed economic asymmetries. Nevertheless, the correlation between connectedness and asset characteristics depends on the choice of test assets. That is, different test assets reflect different patterns in asset pricing, see Feng et al. (2020); Avramov et al. (2025). In this specific exercise, we confirm the prominent status of size as an asset characteristic in building sorted portfolios as test assets (Fama and French, 1993).

Table 9 connects to several literature. For bi-variate portfolios (Panel B), we initially corrobo-

rate Lo and MacKinlay (1990), finding big stocks lead small stocks (coefficient -0.07, row 1) - a result robust to controlling for BM, OP, INV, and MOM (row 2). However, controlling for trading frictions reverses the size coefficient, suggesting big stocks become net receivers, warranting further investigation of size's role in lead-lag effects.[6] Contrary to Chordia and Swaminathan (2000), we find low-turnover stocks transmit signal to high-turnover stocks after controlling for size.[7] It holds for both spread and bi-variate sorts. The divergence from prior papers reflects discretion in test assets and sample periods. Moreover, the two papers focus exclusively on weekly return spillovers, whereas we incorporate multiple firm-level monthly signals, including past returns. Collectively, we demonstrate that cross-asset spillovers are fundamentally linked to asset characteristics.

Figure 3 depicts the TOTAL connectedness index—the average intensity of the off-diagonal elements in Ψ—for both the 138 spread portfolios (dashed line) and the 544 bi-sort portfolios (solid line) over the 1973–2023 period. Four key findings emerge. First, the time-series of TOTAL connectedness on the spread portfolios varies markedly through time: it troughs in the mid-1980s and again after 2020, but peaks around the early 1990s and during the post financial crises, 2010s. Second, the indices for bi-variate sorted portfolios share the trough in mid-1980s and peak in early 1990s, however, slight fluctuations after 2000. Overall, the average level of TOTAL of spread portfolios is almost equal to that of bi-variate sorted portfolios before 2000, but become higher after 2000. Third, despite these episodic surges, the time series reverts to a long-run mean near 0.73, suggesting a stable baseline level of cross-asset information transmission.

Moreover, in the more granular bivariate-sorted universe, we find that the average absolute of diagonal elements of Ψ is 1.34×10^{-3}, while it is 1.32×10^{-3} for the off-diagonal elements—virtually identical in magnitude. This result confirms that both self- and cross-predictive channels are economically significant. Taken together, Figure 3 demonstrates that cross-asset spillover effects intensify during turbulent periods but persist as a pervasive market feature. These findings underscore the importance of modeling the full Ψ matrix—rather than restricting attention to its diagonal elements—for constructing Sharpe ratio-maximizing portfolios.

[6] For comparability, we replicate results for 1973-1987 (matching Lo and MacKinlay (1990)'s sample end) and find consistent size coefficient signs.

[7] While Chordia and Swaminathan (2000) uses "Trading Volume" in their title, they actually employ daily turnover as their volume proxy.

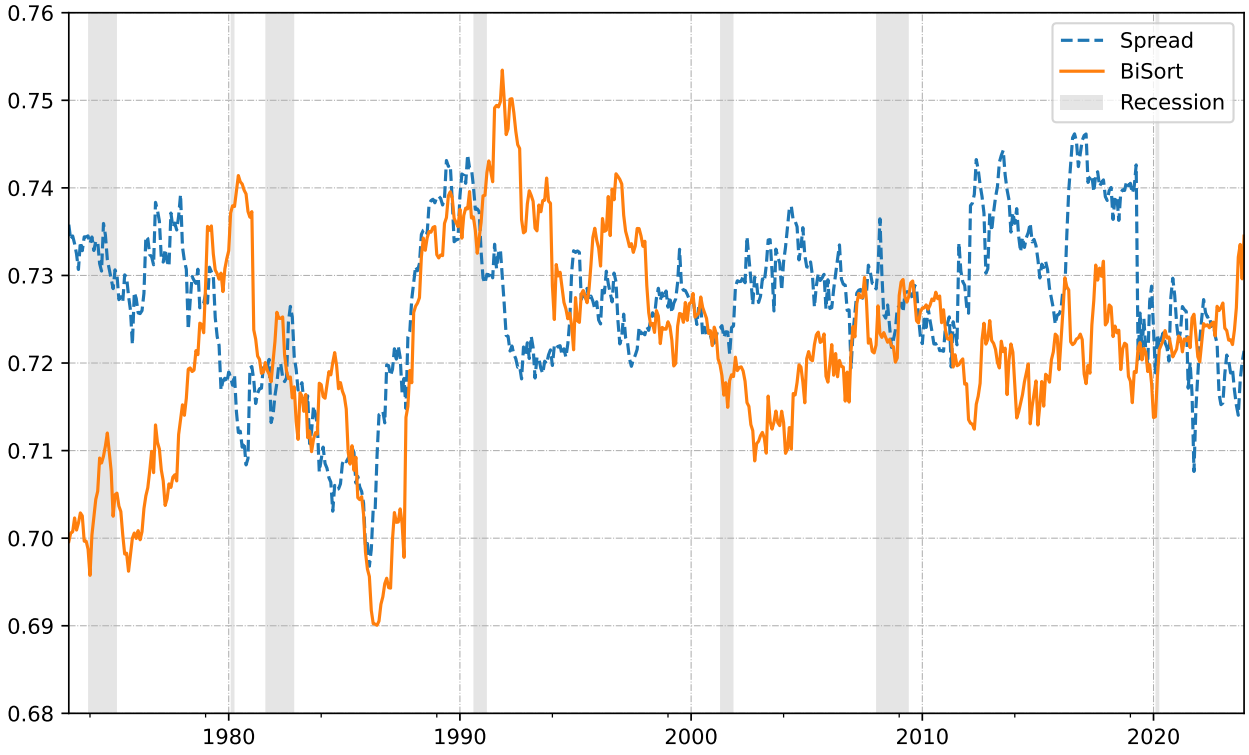
Table 9: Understanding Connectedness

This table reports the time-series average and t-statistics of cross-sectional regressions estimates for each month that regress a connectedness measure on asset characteristics. The assets are 138 spread portfolios. For ease of interpretation, the coefficients are reported with values multiplied by 100. There are three connectedness measures: FROM, TO, and NET. The characteristics of interest are size ("market_equity"), book-to-market equity ("be_me"), operating profits-to-lagged book equity ("ope_bel1"), asset growth ("at_gr1"), price momentum t-12 to t-1 ("ret_12_1"), Amihud illiquidity("ami_126d"), volume("dolvol_126d"), volatility("rvol_21d"), turnover("turnover_126d"), CAPM beta ("beta_60m"), with abbreviations ME, BM, INV, OP, MOM, ILL, VLM, VLT, TRN, and BETA. There is an intercept in the regression, but the estimates are omitted in the table. The sample period is from February 1973 to December 2023.

	ME	BM	OP	INV	MOM	ILL	VLM	VLT	TRN	BETA
					Panel A: Spread Portfolio					
FROM	-0.10									
	(-5.92)									
	-0.20	0.64	0.29	0.02	0.60					
	(-5.86)	(17.93)	(5.95)	(1.17)	(26.45)					
	-0.21	0.75	0.47	-0.01	0.80	-1.09	-0.70	0.47	-0.38	0.06
	(-1.03)	(26.48)	(11.13)	(-0.42)	(28.89)	(-2.47)	(-1.66)	(8.49)	(-4.08)	(1.51)
TO	0.08									
	(2.83)									
	0.18	0.13	-0.20	-0.14	0.22					
	(5.64)	(6.88)	(-7.97)	(-10.13)	(7.78)					
	2.44	0.35	-0.00	-0.17	0.28	1.28	-0.71	0.35	0.11	0.13
	(10.13)	(14.60)	(-0.04)	(-12.28)	(8.57)	(4.02)	(-1.70)	(11.98)	(1.30)	(3.54)
NET	0.17									
	(4.46)									
	0.38	-0.51	-0.48	-0.16	-0.39					
	(7.16)	(-17.23)	(-9.59)	(-9.02)	(-17.21)					
	2.65	-0.40	-0.47	-0.16	-0.53	2.37	-0.02	-0.12	0.49	0.07
	(7.21)	(-13.64)	(-8.32)	(-6.84)	(-15.49)	(4.76)	(-0.03)	(-1.80)	(3.37)	(1.29)
					Panel B: BiSort Portfolio					
FROM	0.01									
	(11.05)									
	0.01	-0.26	-0.09	-0.16	-0.16					
	(2.31)	(-26.81)	(-11.76)	(-26.39)	(-25.21)					
	-0.29	-0.24	-0.08	-0.15	-0.13	0.24	0.66	0.05	-0.25	0.07
	(-8.76)	(-22.91)	(-9.64)	(-28.06)	(-16.39)	(2.23)	(5.75)	(4.03)	(-9.28)	(4.01)
TO	-0.07									
	(-27.65)									
	-0.04	0.05	-0.05	0.04	0.01					
	(-13.16)	(5.70)	(-10.73)	(7.94)	(0.96)					
	0.24	0.07	-0.04	0.02	0.02	0.19	-0.2	0.04	0.23	-0.04
	(10.87)	(6.86)	(-6.66)	(3.70)	(1.42)	(3.83)	(-3.64)	(3.20)	(11.22)	(-4.79)
NET	-0.07									
	(-28.5)									
	-0.05	0.31	0.03	0.20	0.17					
	(-8.75)	(35.6)	(4.11)	(26.24)	(16.59)					
	0.53	0.31	0.04	0.17	0.15	-0.05	-0.85	-0.02	0.48	-0.11
	(11.94)	(31.55)	(4.52)	(21.29)	(12.29)	(-0.40)	(-6.88)	(-1.04)	(12.75)	(-5.56)

Figure 3: Total Connectedness

This figure depicts the time-series plot of total connectedness of Ψ matrix over OOS period form February 1973 to December 2023. The blue dash line is for 138 spread portfolios, and the orange solid line is for 544 bi-variate sorted portfolios. The shadow areas indicate for NBER recession periods.



To analyze directional spillover effects in the bivariate-sorted portfolios more precisely, we decompose the Ψ matrix into four blocks (A, B, C, and D) according to firm size. Figure 4 presents the resulting predictive information flows across these partitions.

Figure 5 presents the time series of absolute average values for each of the four blocks in Ψ. The results reveal consistently stronger predictive relations in Block A (Small $\rightarrow$ Small) and Block C (Big $\rightarrow$ Small) compared to Block B (Small $\rightarrow$ Big) and Block D (Big $\rightarrow$ Big), particularly during the last two decades. The time-series averages measure 1.47 and 1.51 for Blocks A and C, respectively, versus 1.15 and 1.18 for Blocks B and D. Notably, the divergence between the A/C and B/D blocks has increased substantially in recent years.

These findings confirm an asymmetric predictive structure, which aligns with the NET regression coefficient of -0.07 reported in Panel B of Table 9. This result is consistent with the evidence

in Lo and MacKinlay (1990), showing that large stocks tend to lead small stocks, but not vice versa. The persistent and stable nature of these patterns over time supports the economic rationale for imposing restrictions on Ψ, particularly by excluding small-to-large predictive links. Furthermore, the long-run regularity of these asymmetries suggests that dynamic sparsity structures—which adapt to time-varying network block strengths while maintaining economically motivated constraints—could offer significant modeling value.

The overall connectedness analysis reveals that the connection matrix Ψ encodes economically meaningful structure. For bi-variate sorted portfolios on size and other signals, big stocks act as net transmitters of predictive signals; controlling more signals, we find that low trading volume, high turnover ratio, and low-beta stocks are net transmitters. Meanwhile, value, profitable, non-investing, and high-momentum assets are more likely to be net receivers. The strength of cross-predictive relations is comparable to that of self-predictive effects. The overall network intensity fluctuates over time, but remains around a stable level. Decomposing Ψ by firm size shows that predictive flows from large to small firms dominate those in the reverse direction.

Figure 4: Partition of Ψ in Size.

This figure decomposes the Ψ matrix to four blocks based on firm size. They are:

- A: Small (Stock Signals) $\rightarrow$ Small (Stock Returns),
- B: Small $\rightarrow$ Big,
- C: Big $\rightarrow$ Small,
- D: Big $\rightarrow$ Big.

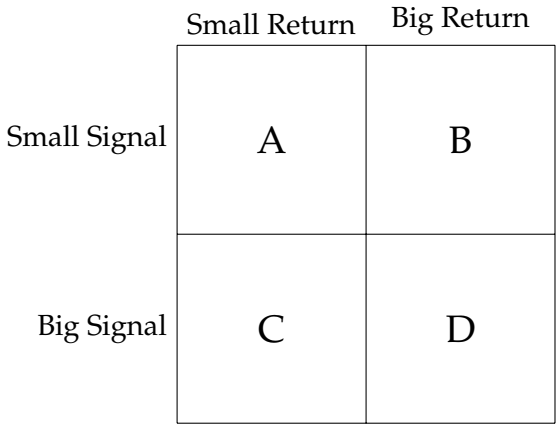
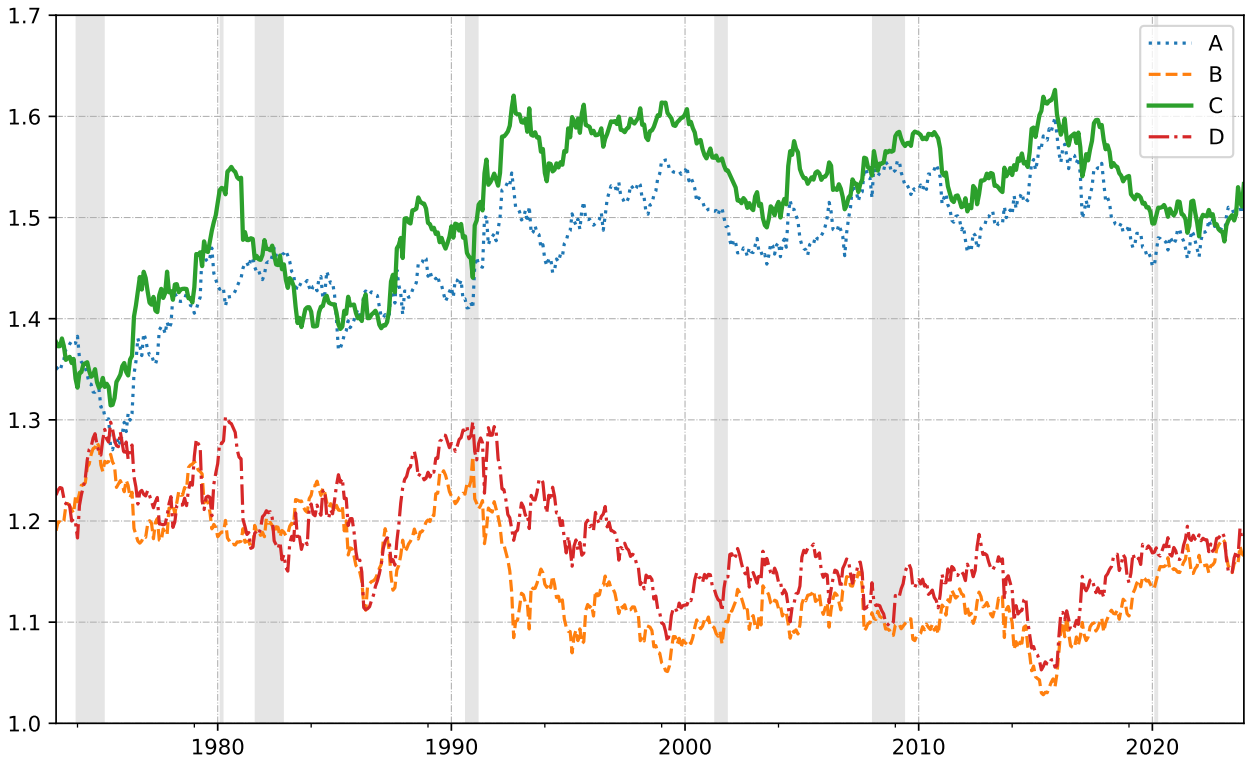


Figure 5: Absolute Average of Four Blocks in Ψ: BiSort Portfolios

This figure shows the time-series plot of the absolute average of elements in four blocks of Ψ. The basic assets are the bi-variate sorted portfolios, where fouce blocks A, B, C, and D represent the strength of cross-predictive relations for (1) small stock signals predict small stock returns, (2) small stock signals predict big stock returns, (3) big stock signals predict small stock returns, and (4) big stock signals predict big stock returns. The sample period is form February 1973 to December 2023. The shadow areas indicate for NBER recession periods.



6 Conclusion

This paper develops a structured framework for constructing Sharpe ratio–maximizing investment strategies using multiple firm-level signals and accounting for informational linkages across assets. By jointly estimating signal relevance and a matrix capturing inter-asset predictive relationships, our approach yields closed-form portfolio weights derived from a generalized eigenvalue decomposition. In high-dimensional settings, estimation is implemented through Ridge-SDF regressions, which offer a stable and interpretable managed-portfolio representation of the decision variables. The resulting stochastic discount factor consistently delivers high out-of-sample Sharpe ratios across a range of asset universes and market conditions, outperforming both self-predictive models and expected return maximization. Economically, the strategy is primarily driven by fundamental characteristics related to investment, valuation, and profitability. In addition, the estimated connection matrix reveals that large and low-turnover stocks tend to act as net transmitters of predictive signals, while the overall strength of inter-asset linkages remains persistently high over time.

Our findings open several promising avenues for future research. First, the framework may be extended to other asset classes where interdependencies among securities are economically plausible, such as corporate bonds, sovereign credit, or international equities. For instance, in the corporate bond market, issuer fundamentals or equity-side information may help forecast bond returns through industry or ownership connections. Second, imposing economically motivated restrictions on the connection matrix—based on supply chain ties, analyst coverage, or institutional co-holdings—could further enhance model interpretability and performance. Finally, applying the framework in a macro-finance context, where predictive relationships span countries or sectors, may yield new insights into the structure of global asset pricing. Together, these avenues offer a foundation for exploring the economic and informational architecture underlying return dynamics across financial markets.

References

Ali, A., L.-S. Hwang, and M. A. Trombley (2003). Arbitrage risk and the book-to-market anomaly. *Journal of Financial Economics 69*(2), 355–373.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets 5*(1), 31–56.

Avramov, D., S. Cheng, and L. Metzker (2023). Machine learning vs. economic restrictions: Evidence from stock return predictability. *Management Science 69*(5), 2587–2619.

Avramov, D., G. Feng, J. He, and S. Xiao (2025). Are asset pricing models sparse? Technical report, Reichman University (IDC), Herzliya, Israel.

Back, K. (2017). *Asset pricing and portfolio choice theory: 2nd edition*. Oxford University Press.

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

Belo, F., X. Lin, and S. Bazdresch (2014). Labor hiring, investment, and stock return predictability in the cross section. *Journal of Political Economy 122*(1), 129–177.

Britten-Jones, M. (1999). The sampling error in estimates of mean-variance efficient portfolio weights. *Journal of Finance 54*(2), 655–671.

Chen, L., M. Pelger, and J. Zhu (2024). Deep learning in asset pricing. *Management Science 70*(2), 714–750.

Chordia, T. and B. Swaminathan (2000). Trading volume and cross-autocorrelations in stock returns. *The Journal of Finance 55*(2), 913–935.

Cochrane, J. H. (2009). *Asset pricing: Revised edition*. Princeton university press.

Cochrane, J. H. (2011). Presidential address: Discount rates. *Journal of finance 66*(4), 1047–1108.

Cohen, L. and A. Frazzini (2008). Economic links and predictable returns. *Journal of Finance 63*(4), 1977–2011.

Cohen, L. and D. Lou (2012). Complicated firms. *Journal of Financial Economics 104*(2), 383–400.

Cong, L. W., G. Feng, J. He, and X. He (2025). Growing the efficient frontier on panel trees. *Journal of Financial Economics 167*, 104024.

Cong, L. W., K. Tang, J. Wang, and Y. Zhang (2022). Alphaportfolio: Direct construction through deep reinforcement learning and interpretable ai. Technical report, Cornell University.

Cooper, M. J., H. Gulen, and M. J. Schill (2008). Asset growth and the cross-section of stock returns. *Journal of Finance 63*(4), 1609–1651.

Corwin, S. A. and P. Schultz (2012). A simple way to estimate bid-ask spreads from daily high and low prices. *Journal of Finance 67*(2), 719–760.

Daniel, K., D. Hirshleifer, and L. Sun (2020). Short-and long-horizon behavioral factors. *Review of Financial Studies 33*(4), 1673–1736.

Dechow, P. M., R. G. Sloan, and M. T. Soliman (2004). Implied equity duration: A new measure of equity risk. *Review of Accounting Studies 9*, 197–228.

Dichev, I. D. (1998). Is the risk of bankruptcy a systematic risk? *Journal of Finance 53*(3), 1131–1147.

Didisheim, A., S. B. Ke, B. T. Kelly, and S. Malamud (2024). Apt or "aipt"? the surprising dominance of large factor models. Technical report, Yale University.

Diebold, F. X. and K. Yılmaz (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics 182*(1), 119–134.

Eugene, F. and K. French (1992). The cross-section of expected stock returns. *Journal of Finance 47*(2), 427–465.

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

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

Feng, G., S. Giglio, and D. Xiu (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance 75*(3), 1327–1370.

Feng, G., J. He, N. G. Polson, and J. Xu (2024). Deep learning in characteristics-sorted factor models. *Journal of Financial and Quantitative Analysis 59*(7), 3001–3036.

Green, J., J. R. Hand, and X. F. Zhang (2017). The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies 30*(12), 4389–4436.

Gu, S., B. Kelly, and D. Xiu (2021). Autoencoder asset pricing models. *Journal of Econometrics 222*(1), 429–450.

Harvey, C. R., Y. Liu, and H. Zhu (2016). ... and the cross-section of expected returns. *Review of Financial Studies 29*(1), 5–68.

Haugen, R. A. and N. L. Baker (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics 41*(3), 401–439.

He, S., M. Yuan, and G. Zhou (2024). Principal portfolios: The multi-signal case. Technical report, Washington University in St. Louis.

Hou, K. (2007). Industry information diffusion and the lead-lag effect in stock returns. *The review of financial studies 20*(4), 1113–1138.

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

Jensen, T. I., B. Kelly, and L. H. Pedersen (2023). Is there a replication crisis in finance? *Journal of Finance 78*(5), 2465–2518.

Kelly, B., B. Kuznetsov, S. Malamud, and T. A. Xu (2024). Artificial intelligence asset pricing models. Technical report, Yale University.

Kelly, B., S. Malamud, and L. H. Pedersen (2023). Principal portfolios. *Journal of Finance 78*(1), 347–387.

Kelly, B. and D. Xiu (2023). Financial machine learning. *Foundations and Trends® in Finance 13*(3-4), 205–363.

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

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Lakonishok, J., A. Shleifer, and R. W. Vishny (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance 49*(5), 1541–1578.

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

Litzenberger, R. H. and K. Ramaswamy (1979). The effect of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of financial economics 7*(2), 163–195.

Liu, Y., G. Zhou, and Y. Zhu (2025). Maximizing the sharpe ratio: A genetic programming approach. Technical report, Washington University in St. Louis.

Lo, A. W. and A. C. MacKinlay (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies 3*(2), 175–205.

Luo, B., G. Zhou, and T. Zhou (2025). Which factors matter in the pricing kernel? Technical report, Washington University in St. Louis.

Miller, M. H. and M. S. Scholes (1982). Dividends and taxes: Some empirical evidence. *Journal of Political Economy 90*(6), 1118–1141.

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

Ortiz-Molina, H. and G. M. Phillips (2014). Real asset illiquidity and the cost of capital. *Journal of Financial and Quantitative Analysis 49*(1), 1–32.

Pastor, L. and R. F. Stambaugh (2003). Liquidity risk and expected stock returns. *Journal of Political Economy 111*(3), 642–685.

Richardson, S. A., R. G. Sloan, M. T. Soliman, and I. Tuna (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics 39*(3), 437–485.

Shen, Z. and D. Xiu (2024). Can machines learn weak signals? Technical report, University of Chicago.

Soliman, M. T. (2008). The use of dupont analysis by market participants. *The Accounting Review 83*(3), 823–853.

Stambaugh, R. F. and Y. Yuan (2017). Mispricing factors. *Review of Financial Studies 30*(4), 1270–1315.

Appendices

A Proof of Proposition 1

Expected Return We first express π_s, the realized return on the trading strategy. Recognize that $\pi_s = \Lambda' S_t' \Psi r_s = \sum_{i=1}^{N} \Psi_i' r_s S_{it}' \Lambda = \text{tr}\left[\Lambda \sum_{i=1}^{N} \Psi_i' \Pi_{si}\right] = \Lambda' \Pi_s' \Phi$, where Ψ_i' is a $1 \times N$ vector which is the i-th row of Ψ, S_{it}' is a $1 \times M$ vector, which is the i-th row of S_t, tr stands for the trace operator, and Π_s is an $N^2 \times M$ matrix that vertically stacks the $N \times M$ matrices $\Pi_{si} = r_s S_{it}'$ for $i = 1, 2, \cdots, N$.

Then, on the basis of realized return, the expected value is given by

$$E(\pi_s) = \Lambda' \Pi' \Phi = \Phi' \Pi \Lambda. \tag{A.1}$$

Variance Let Σ_Φ be the covariance matrix of $\text{vec}(\Pi_s')$. We express π_s in terms of $\text{vec}(\Pi_s')$:

$$\pi_s = \Lambda' \Pi_s' \Phi = \Lambda' \text{vec}(\Pi_s' \Phi). \tag{A.2}$$

Using the property of vectorization:

$$\text{vec}(ABC) = (C' \otimes A)\text{vec}(B) \tag{A.3}$$

we get:

$$\text{vec}(\Pi_s' \Phi) = (\Phi' \otimes I_M)\text{vec}(\Pi_s') \tag{A.4}$$

Therefore:

$$\pi_s = \Lambda'(\Phi' \otimes I_M)\text{vec}(\Pi_s') \tag{A.5}$$

The variance of π_s is:

$$\begin{aligned} \text{Var}(\pi_s) &= \Lambda'(\Phi' \otimes I_M)\Sigma_\Phi(\Phi \otimes I_M)\Lambda, \\ &= \Lambda' B_\Phi \Lambda, \end{aligned} \tag{A.6}$$

where $B_\Phi = (\Phi' \otimes I_M)\Sigma_\Phi(\Phi \otimes I_M)$.

We consider an alternative expression of Var(π_s). Let Σ_Λ be the covariance matrix of vec(Π_s). We express π_s in terms of vec(Π_s):

$$\pi_s = \Phi' \Pi_s \Lambda = \Phi' \text{vec}(\Pi_s \Lambda). \tag{A.7}$$

Again using the property of vectorization, we get:

$$\text{vec}(\Pi_s \Lambda) = (\Lambda' \otimes I_{N^2}) \text{vec}(\Pi_s). \tag{A.8}$$

Therefore:

$$\pi_s = \Phi' (\Lambda' \otimes I_{N^2}) \text{vec}(\Pi_s) \tag{A.9}$$

The variance of π_s is:

$$\begin{aligned} \text{Var}(\pi_s) &= \Phi' (\Lambda' \otimes I_{N^2}) \Sigma_\Lambda (\Lambda \otimes I_{N^2}) \Phi, \\ &= \Phi' B_\Lambda \Phi, \end{aligned} \tag{A.10}$$

where $B_\Lambda = (\Lambda' \otimes I_{N^2}) \Sigma_\Lambda (\Lambda \otimes I_{N^2})$.

Sharpe Ratio With the expected return and variance, we express the Sharpe ratio square as:

$$SR^2 = \frac{\Lambda' A_\Phi \Lambda}{\Lambda' B_\Phi \Lambda}, \tag{A.11}$$

where $A_\Phi = \Pi' \Phi \Phi' \Pi, B_\Phi = (\Phi' \otimes I_M) \Sigma_\Phi (\Phi \otimes I_M)$, and Σ_Φ is the covariance matrix of vec(Π_s'). Alternatively, we express the Sharpe ratio squared as:

$$SR^2 = \frac{\Phi' A_\Lambda \Phi}{\Phi' B_\Lambda \Phi}, \tag{A.12}$$

where $A_\Lambda = \Pi \Lambda \Lambda' \Pi', B_\Lambda = (\Lambda' \otimes I_{N^2}) \Sigma_\Lambda (\Lambda \otimes I_{N^2})$, and Σ_Λ is the covariance matrix of vec(Π_s). These alternative expressions of SR^2 assist in finding the solution to maximize the Sharpe ratio, with details in Appendix E.

B Proof of Expected Return Reduction due to Zero-Cost Constraint

Consider the matrix Π formed by vertically stacking N matrices Π_i, each of dimension $N \times M$, and let $\tilde{\Pi}$ be the matrix obtained after pre-multiplying each Π_i by the matrix Θ, where $\Theta = I_N - \frac{1}{N}\iota_N \iota_N'$. Here, Θ is a projection matrix that projects vectors onto the space orthogonal to the vector ι_N of ones.

Properties of Θ:

- Θ is symmetric and idempotent, i.e., $\Theta^2 = \Theta$ and $\Theta' = \Theta$, confirming that it is a projection matrix.
- The eigenvalues of Θ are 0 along the direction of ι_N and 1 along all directions orthogonal to ι_N.

Impact on Singular Values:

1. The matrix Θ modifies Π_i by removing its component in the direction of ι_N. This operation reduces the variance in Π_i that is aligned with ι_N.
2. Given the singular value decomposition of $\Pi = U\Sigma V'$, the transformation $\tilde{\Pi} = (\Theta\Pi_i)$ can be viewed through the lens of modified singular vectors. Since Θ acts as an identity on the space orthogonal to ι_N and zeroes out components along ι_N, it does not increase the magnitude of any singular vector components.
3. The singular values $\lambda_i(\tilde{\Pi})$ of the transformed matrix $\tilde{\Pi}$ correspond to the norms of the vectors ΘU_i, where U_i are the left singular vectors of Π. Since Θ is a projection (and thus a norm-reducing operation except where it acts as the identity), we have:

$$\|\Theta U_i\| \leq \|U_i\|$$

4. Therefore, the singular values of $\tilde{\Pi}$ must satisfy:

$$\lambda_i(\tilde{\Pi}) \leq \lambda_i(\Pi)$$

for each i, because the projection does not increase vector norms and reduces them for vectors with non-zero components in the direction of ι_N.

To be more precise, the highest singular value of the transformed matrix does not change due to the preservation of the highest singular value by Θ. However, the transformation induced by Θ results in a reduction of singular values in the transformed matrix $\tilde{\Pi}$ in the other singular values, leading to a decrease in variance explained by certain components. Specifically, at least one singular value of $\tilde{\Pi}$ is strongly diminished compared to the corresponding singular value of the original matrix Π. This reduction underscores the effectiveness of the transformation in diminishing the influence of certain components in Π and highlights its role in variance reduction. Hence, both expected return and risk of the trading strategy are lower in the presence of the zero-cost restriction.

C Proof and Derivations for Propositions 4 and 5

This section focuses on maximizing the squared Sharpe ratio of a linear strategy. The results extend naturally to the Sharpe ratio maximization of a nonlinear strategy with an augmented signal space.

Maximizing the squared Sharpe ratio constitutes a *generalized Rayleigh quotient problem*, which can be solved via an *eigenvalue problem*. However, in empirical settings, the solution to this *eigenvalue problem* often becomes ill-conditioned in high-dimensional settings.

To address this issue, we employ Ridge-SDF regressions to estimate the decision variables, providing an intuitive managed-portfolio interpretation. Finally, we present an iterative algorithm to estimate Λ and Φ until convergence. The details are as follows.

Define the squared Sharpe ratio as a function of Λ. According to Proposition 1, the squared Sharpe ratio takes the form:

$$SR^2 = \frac{\Lambda' A_\Phi \Lambda}{\Lambda' B_\Phi \Lambda}, \tag{C.13}$$

where $A_\Phi = \Pi'\Phi\Phi'\Pi$, $B_\Phi = (\Phi' \otimes I_M)\Sigma_\Phi(\Phi \otimes I_M)$, and Σ_Φ is the covariance matrix of $\text{vec}(\Pi_s')$.

Maximizing the squared Sharpe ratio with respect to Λ. From (A.11), the optimization problem is formulated as:

$$\max_{\Lambda} \frac{\Lambda' A_\Phi \Lambda}{\Lambda' B_\Phi \Lambda}. \tag{C.14}$$

This is equivalent to the constrained optimization problem:

$$\max_{\Lambda} \Lambda' A_\Phi \Lambda \quad \text{s.t.} \quad \Lambda' B_\Phi \Lambda = \kappa. \tag{C.15}$$

Given the norm constraint on Λ, we set $\kappa = 1$ without loss of generality.

Applying the method of Lagrange multipliers, we define the Lagrangian function:

$$L(\Lambda, \lambda) = \Lambda' A_\Phi \Lambda - \lambda(\Lambda' B_\Phi \Lambda - 1). \tag{C.16}$$

Taking derivatives with respect to Λ yields the *generalized eigenvalue problem*:

$$A_\Phi \Lambda = \lambda B_\Phi \Lambda. \tag{C.17}$$

Multiplying both sides by B_Φ^{-1} gives:

$$B_\Phi^{-1} A_\Phi \Lambda = \lambda \Lambda. \tag{C.18}$$

Defining $C_\Phi = B_\Phi^{-1} A_\Phi$, we obtain the standard *eigenvalue problem*:

$$C_\Phi \Lambda = \lambda \Lambda. \tag{C.19}$$

Solving (C.19) provides the eigenvector corresponding to the largest eigenvalue, Λ_{max}. Normalizing for the norm constraint, we set:

$$\Lambda = \frac{\Lambda_{\text{max}}}{||\Lambda_{\text{max}}||}. \tag{C.20}$$

Since the solution for Λ depends on Φ, we define the function:

$$\Lambda = \arg\max_{\Lambda} \frac{\Lambda' A_\Phi \Lambda}{\Lambda' B_\Phi \Lambda} = \mathbf{\Lambda}(\Phi). \tag{C.21}$$

Estimating high-dimensional Λ using ridge regression. In high-dimensional settings where M is large relative to the number of observations T, the solution in (C.21) often fails in out-of-sample tests.

To address this, consider a set of managed portfolios χ_Φ of dimension $T \times M$:

$$\chi_\Phi = \Pi' \Phi. \tag{C.22}$$

The optimization in (C.21) is an asset allocation problem in which the goal is to determine the investment weights Λ for the managed portfolios χ_Φ to maximize the squared Sharpe ratio. This is equivalent to estimating Λ as the mean-variance efficient portfolio weights.

Following Britten-Jones (1999), we estimate Λ using the regression:

$$\mathbf{1} = \chi_\Phi \Lambda + \mathbf{u}, \tag{C.23}$$

where $\mathbf{1}$ is a T-vector of ones.

To improve out-of-sample performance, we adopt ridge regression, as in Shen and Xiu (2024); Didisheim et al. (2024):

$$\hat{\Lambda} = (\chi_\Phi' \chi_\Phi + \lambda I_M)^{-1} \chi_\Phi' \mathbf{1}, \tag{C.24}$$

where λ is a shrinkage parameter. The solution in (C.24) coincides with (C.21) when $\lambda = 0$. While (C.24) may underperform in-sample, it improves robustness for out-of-sample applications.

Define the squared Sharpe ratio as a function of Φ. Alternatively, we express the squared Sharpe ratio as:

$$SR^2 = \frac{\Phi' A_\Lambda \Phi}{\Phi' B_\Lambda \Phi}, \tag{C.25}$$

where $A_\Lambda = \Pi \Lambda \Lambda' \Pi'$, $B_\Lambda = (\Lambda' \otimes I_{N^2}) \Sigma_\Lambda (\Lambda \otimes I_{N^2})$, and Σ_Λ is the covariance matrix of $\text{vec}(\Pi_s)$.

Maximizing the squared Sharpe ratio with respect to Φ. Referring to Eq. (A.12), we formulate the optimization problem as:

$$\Phi = \arg\max_\Phi \frac{\Phi' A_\Lambda \Phi}{\Phi' B_\Lambda \Phi} = \mathbf{\Phi}(\Lambda). \tag{C.26}$$

Solving (C.26) follows the same procedure as (C.21).

Estimating high-dimensional Φ using ridge regression. Analogous to (C.24), we define managed portfolios χ_Λ of dimension $T \times N^2$:

$$\chi_\Lambda = \Pi\Lambda. \tag{C.27}$$

Applying ridge regression, we estimate Φ as:

$$\hat{\Phi} = (\chi_\Lambda'\chi_\Lambda + \lambda I_{N^2})^{-1}\chi_\Lambda'\mathbf{1}. \tag{C.28}$$

Algorithm and Iteration. To solve the whole problem, we do iterations until convergence. In each iteration, we have four steps:

1. Given $\Pi, \Sigma_\Lambda, \Lambda$, update the values of $A_\Lambda, B_\Lambda, C_\Lambda$,
2. Solve Eq.(C.28) to get the updated Φ,
3. Given Π, Σ_Φ, Φ, update the values of A_Φ, B_Φ, C_Φ,
4. Solve Eq.(C.24) to get the updated Λ.

A full description of the algorithm is in Algorithm 1.

Algorithm 1 Maximize Sharpe Ratio

procedure MAXSR(Λ, Φ)
Input Asset returns r_s and signals S_t.
outcome Investment decision variables Λ, Φ.
    Calculate $\Pi, \Sigma_\Phi, \Sigma_\Lambda$. ▷ These variables are Constant.
    Initialize index of iteration $k = 0$. ▷ We use k in notation $\Lambda^{\{k\}}, \Phi^{\{k\}}$.
    Initialize $\Lambda^{\{0\}}$. ▷ E.g., the solution in Max Expected Return strategy.
    while Termination Conditions *not* Activated **do**
        Update $A_\Lambda, B_\Lambda, C_\Lambda$ with $\Lambda^{\{k\}}$.
        Update $\Phi^{\{k+1\}}$ by solving Eq.(23).

$$\Phi^{\{k+1\}} = \mathbf{\Phi}(\Lambda^{\{k\}}).$$

        Update A_Φ, B_Φ, C_Φ with $\Phi^{\{k+1\}}$,
        Update $\Lambda^{\{k+1\}}$ by solving Eq.(21).

$$\Lambda^{\{k+1\}} = \mathbf{\Lambda}(\Phi^{\{k+1\}}).$$

        $k = k + 1$.
    end while
    return $\Lambda^{\{k\}}, \Phi^{\{k\}}$
end procedure

D Robustness of λ the Ridge Shrinkage Parameter

Table D.1 reports the performance of Sharpe-maximizing strategy under an array of values for λ, the single hyperparameter in ridge SDF regression. Recall that the default setting in the paper is $\lambda = 10^{-4}$. For other values around the default, the OOS Sharpe ratios are stable. For instance, in Panel A: Spread Portfolio, all the Sharpe ratios are above 2.0; in Panel B: BiSort Portfolios, all the Sharpe ratios are above 3.5.

Table D.1: Robustness of the λ shrinkage in Ridge SDF Regression

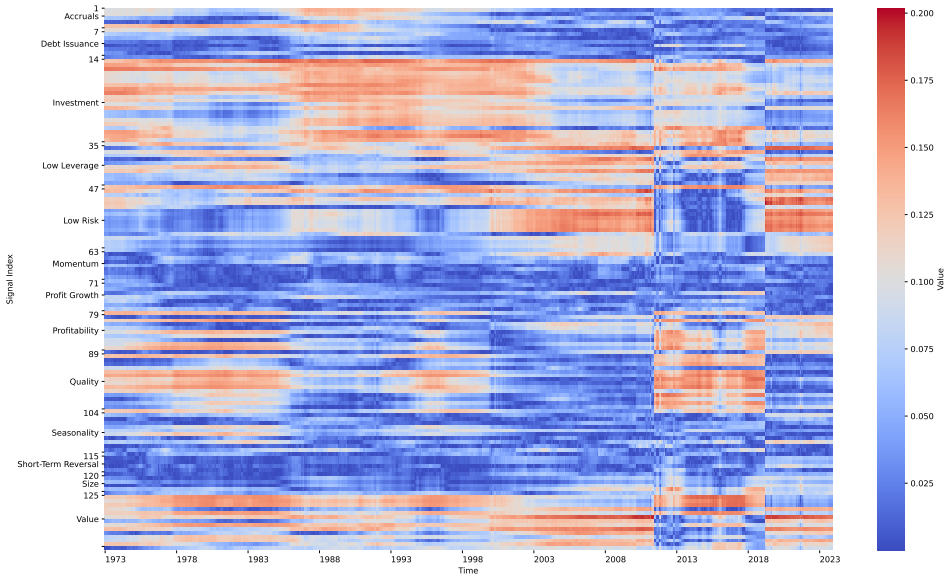
This table reports the properties of the Sharpe-maximizing strategies with respect to the shrinkage parameter λ in ridge regressions (21) and (23). The first column shows the value of the shrinkage parameter, including $10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$, where 10^{-4} is the default setting in main results. The remaining columns report the properties include monthly average returns (%), monthly standard deviation(%), and annualized Sharpe ratio, time-series average of the sum of positions on basic assets, and time-series average of the absolute sum of positions on basic assets.

Shrinkage	Avg%	Std%	SR	Sum	ASum
		Panel A: Spread Portfolios			
10^{-2}	0.375	0.560	2.320	0.00	2.00
10^{-3}	0.246	0.352	2.422	0.00	2.00
10^{-4}	0.175	0.278	2.191	0.00	2.00
10^{-5}	0.155	0.262	2.050	0.00	2.00
10^{-6}	0.151	0.259	2.019	0.00	2.00
		Panel B: BiSort Portfolios			
10^{-2}	0.353	0.331	3.700	0.00	2.00
10^{-3}	0.226	0.209	3.751	0.00	2.00
10^{-4}	0.183	0.175	3.622	0.00	2.00
10^{-5}	0.175	0.168	3.593	0.00	2.00
10^{-6}	0.173	0.167	3.593	0.00	2.00

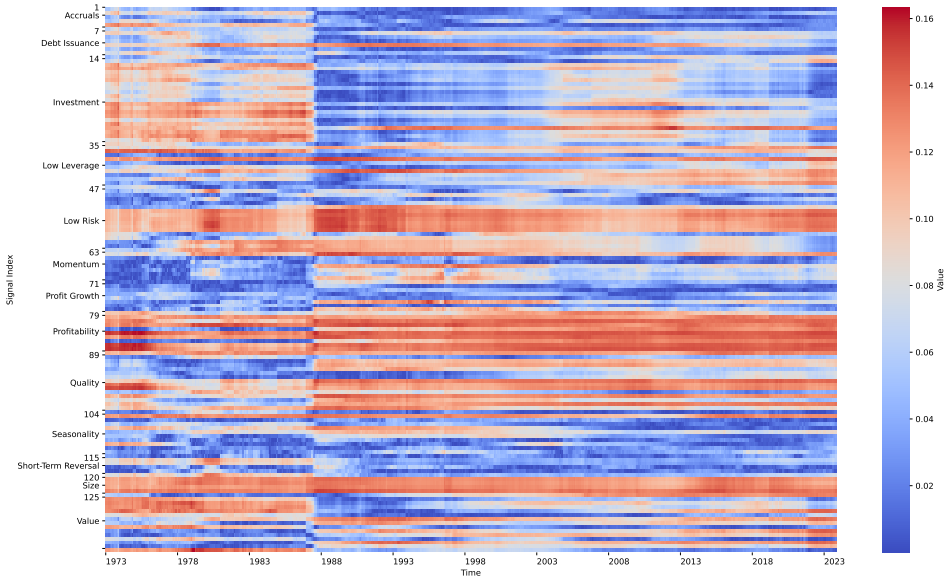
E Signal-Level Importance

Figure E.1: Signal Importance

This figure complements the theme-level signal importance in Figure 2 by providing the 138 signal-level importance in full detail. These signals of the same theme are grouped in the vertical axis, where the 13 themes follow Jensen et al. (2023). For interpretation, we focus on the absolute value of elements in Λ. Sub-figures a and b report for spread portfolios and bi-variate sorted portfolios, respectively.



(A) Spread Portfolio



(B) BiSort Portfolio