Matteo Bagnara

The Economic Value of Cross-Predictability: A Performance-Based Measure

SAFE Working Paper No. 424 | July 2024

The Economic Value of Cross-predictability: A Performance-based Measure

MATTEO BAGNARA*

ABSTRACT

Cross-predictability denotes the fact that some assets can predict other assets' returns. I propose a novel performance-based measure that disentangles the economic value of cross-predictability into two components: the predictive power of one asset's signal for other assets' returns (*cross-predictive signals*) and the amount of an asset's return explained by other assets' signals (*cross-predicted returns*). Empirically, the latter component dominates the former in the overall cross-prediction effects. In the cross-section, cross-predictability gravitates towards small firms that are strongly mispriced and difficult to arbitrage, while it becomes more difficult to cross-predict returns when market capitalization and book-to-market ratio rise.

JEL classification: C58, G11, G12, G14

Keywords: Empirical Asset Pricing, Portfolio Choice, Expected Returns, Cross-Predictability

*Leibniz Institute for Financial Research SAFE, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 3, 60323, Frankfurt am Main, Germany, bagnara@safe-frankfurt.de.
I thank Christian Schlag for useful comments. I gratefully acknowledge research support from the Leibniz Institute for Financial Research SAFE.

1 Introduction

"Firms do not exist as independent entities, but are linked to each other through many types of relationships. Some of these links are clear and contractual, while others are implicit and less transparent."

Cohen and Frazzini (2008)

Economic links among different firms can give rise to cross-predictability and profit opportunities. For instance, a long-short trading strategy exploiting customer-supplier relations earns monthly alphas of over 150 basis points over the period 1980-2004 (Cohen and Frazzini, 2008). Lead-lag relations between large and small stocks contribute to more than 50% of profits from contrarian strategies (Lo and MacKinlay, 1990), which means they can help explaining important phenomena occurring in financial markets. Nevertheless, standard asset pricing almost exclusively focuses on own-asset predictive signals for building characteristic-sorted portfolios or for pricing tests in the spirit of Gibbons et al. (1989) (Kelly et al., 2023). In other words, the predictive content of signals coming from other assets is typically ignored when looking for anomalies or new predictors. Cross-predictability is often overlooked likely due to the lack of a solid understanding of its origin. If, on the one hand, a few conjectures have been proposed, such as investors' limited attention or market segmentation, on the other hand a unifying framework has yet to be found that is consistent with all cross-prediction patterns that have been documented.

In this paper, I provide a novel approach to quantify the economic value of cross-predictability with a performance-based measure that builds on the recent work of Kelly et al. (2023) (hereafter KMP). This method allows to identify the sources of cross-prediction effects in the cross-section and to distinguish two components: *cross-predictive signals*, i.e., the power that one asset's signal has to predict other assets' returns; and *cross-predicted returns*, i.e., the amount of one asset's return that is explained by other assets' signals. I find that neglecting to predict returns of small stocks subject to limits to arbitrage using other assets' signals results in dramatic losses in portfolio performance, whereas it becomes increasingly more difficult to cross-predict returns when market capitalization and book-to-market ratio rise.

In the literature, cross-predictability is often interpreted as the existence of lead-lag relations, where the returns of a group of stocks predict the returns of another group of stocks in the following period. Economic links leading to these effects are found between large and small stocks, firms in customer-supply relationships (Menzly and Ozbas, 2010), stocks with a higher and lower analyst coverage (Brennan et al., 1993) and institutional ownership (Badrinath et al., 1995), or that are more and less frequently traded (Chordia and Swaminathan, 2000). The approach used so far to measure cross-predictability mainly consists in constructing a long-short portfolio obtained using binary splits of the cross-section that follows one of these economic links. For example, buying stocks of firms whose customer performed well in the previous period reveals a profitable strategy (Cohen and Frazzini, 2008). As other forms of return predictability, lead-lag effects go against the notion of market efficiency, and justifications such as slow information diffusion, investors' specialization, or even information manifestation (Huang et al., 2022), are needed. Alternatively, firms in leading industries might provide resolution of uncertainty relative to companies in lagging industries (Croce et al., 2023). However, "[...] cross-predictability effects are different from lead-lag effects where one market either always leads or always lags another market [...] With cross-predictability, a market can sometimes lead and sometimes lag another related market depending on where the information originates." (Menzly and Ozbas, 2010). Put differently, cross-predictability effects are likely complex, and limiting the analysis to coarse partitions of the cross-section (such as customer-supplier splits) is ill-suited to provide further insights on the origin of cross-predictability and measure its economic value. This is where Principal Portfolio Analysis (PPA, Kelly et al. (2023)) comes in especially handy.

PPA solves a general portfolio optimization problem using linear strategies that encompass other well-known strategies such as long-short portfolios (e.g. Fama and French (1993)). The solution delivered by this method are Principal Portfolios (PPs), which are ordered by expected return and can be broken down into an alpha and a beta component. PPs prove very profitable, beating both the long-short factor based on the same firm signal and a simple factor that trade all assets according to the strength of only their own signal. The driver of this outperformance lies in a simple mathematical object, the *prediction matrix*, which contains information not only about the predictive power of firm characteristics for their own return, but also for other assets' returns. Accounting for cross-predictability effects allows to improve portfolio performance meaningfully.

By setting specific restrictions on the prediction matrix, I are able to quantify how much cross-predictability contributes to PPs' performance, distinguishing between how much one signal cross-predicts other assets' returns from how much one asset's returns are cross-predicted by other assets' signals. Measuring cross-predictability through PPA represents a noticeable improvement compared to existing methods that focus on simple long-short portfolios, because PPs result from an explicit optimization process where all securities' signals predict each individual return. In this framework I are able to control for any other effect that could otherwise influence the result of lead-lag strategies, leading to a clear-cut identification of the sources of economic value contained in cross-prediction effects among assets. Furthermore, using PPA permits capturing cross-predictability relative to any firm characteristic and not just past returns, in contrast to previous studies.

The findings are three-fold. First, the economic value of cross-predictability concentrates on small stocks. The average realized return and the Sharpe Ratio (SR) of the first PP built on ten size-sorted portfolios decrease by roughly 60% and 50%, respectively, if I exclude any type of cross-prediction effects involving the smallest size-decile. Moreover, while there is valuable information in predicting returns of small stocks using signals from other assets, small firms' signals contain low information to predict the rest of the cross-section. Imposing no prediction from non-small stocks' characteristics to small stocks' returns reduces both the average return and the SR of the first PP by around 25%. Excluding small companies' cross-predictive signals barely changes the portfolio performance.

Second, the larger the market capitalization, the more difficult it is to predict stocks from other assets' signals, up to the point of becoming detrimental. Excluding cross-prediction effects for each size decile has an impact on the PP performance that decreases to zero around the fifth decile, and becomes slightly negative afterwards. Said differently, it is beneficial to avoid predicting larger stocks with other assets' signals for investment purposes. On the other hand, the cross-predictive power of an asset rises with size. The middle of the cross-section (i.e., mid-cap stocks) has instead a negligible impact. The dominant role of small firms drives also the composition of PPs, which are strongly tilted towards these assets.

Third, cross-predictability is higher for stocks more prone to mispricing due to limits to arbitrage. Cross-sectional regressions show that idiosyncratic volatility and mispricing measures are strongly positively associated with cross-predictability measures, while size, book-to-market and operating profitability exert a negative influence. These characteristics have opposite relations to cross-predictive signals, thereby suggesting that limits to arbitrage provide room not only for own-asset predictability but also for cross-predictability, at the same time decreasing the strength of one asset's signal to predict other assets' returns. These results are robust to controlling for value nonparametrically through double-sorted portfolios.

The rest of the paper is organized as follows. Section 2 reviews the literature, highlighting my contributions. Section 3 provides a primer on KMP's PPA and benchmarks it with the construction of risk factors typical in empirical asset pricing. Section 4 explains the methodology employed in this paper to measure cross-predictability. Section 5 illustrates the data used. Section 6 presents the main results. Section 7 proposes an explanation for the findings. Section 8 discusses some robustness tests. Section 9 concludes.

2 Relation to the Literature and Contribution

Starting from the pioneering works of Lo and MacKinlay (1990), cross-predictability has been mainly studied through the lens of lead-lag relations between groups of firms linked by economic relations. Besides the studies mentioned above, more recent contributions observe cross-predictability based on size within the same industry (Hou, 2007), conglomerates (Cohen and Lou, 2012), technological similarity (Lee et al., 2019), geographic position (Parsons et al., 2020), shared analyst coverage (Ali and Hirshleifer, 2020) and production network (Gofman et al., 2020). These studies focus on lead-lag relations between just two groups of firms, such as large and small stocks. Instead, I look at potential cross-prediction effects along the entire cross-section by means of finer partitions based on the distribution of salient firm characteristics like size and book-to-market ratio.

Notably, cross-predictability has been documented also at the industry level (Schlag and Zeng, 2019; Croce et al., 2023) and in international markets (Rapach et al., 2013), which shows that it is a phenomenon that has a strong impact on asset prices.

Up until now, only few studies leverage information contained in assets' characteristics to make cross-predictions. Kelly and Pruitt (2013) use the a cross-sectional measure of book-to-market ratio to predict returns of both aggregate market, characteristic-based portfolios and industries. In a similar spirit, Detzel and Strauss (2018) show that the cross-section of book-to-market beyond a market-level factor helps forecasting portfolio returns. Müller (2019) identifies economic links among firms through well-known characteristics and builds a measure of earning surprise across different signals that predicts individual stock returns. As for the literature related to lead-lag relations, these papers take into account only the extremes of firm characteristics distributions to build trading strategies that fundamentally hinge on long-short portfolios, disregarding what happens in the middle. Said differently, I have no idea whether cross-prediction effects exists from the extremes towards the middle or vice-versa. I fill this gap by employing a method that allows connections among all the parts of the cross-section without ruling out any *a priori*. The approach is flexible enough that it can control for the effect of a specific signal by simple means of double-sorted portfolios. Furthermore, I provide a straightforward performance-based measure to quantify the economic importance of cross-prediction effects in the context of portfolio optimization. To the best of my knowledge, this is the first work achieving this result.

my work is related to the growing literature that employs machine learning in asset pricing. Prominent examples are Gu et al. (2020), Kelly et al. (2019), Kozak et al. (2020), Freyberger et al. (2020), Goodarzi et al. (2022).[1] Among others, Kozak et al. (2018) and Lettau and Pelger (2020) evaluate their pricing factors using only the top- and bottom-decile portfolios (or on long-short portfolios) as "most of the relevant information is contained in the extreme first and tenth decile portfolios." (p.2292). Here I provide the evidence that has been missing to sustain this argument, confirming the previous intuition in the literature.

3 Principal Portfolios: A Primer

This section provides a framework encompassing both traditional trading strategies based on long-short portfolios, on own-asset signals only, and the PPs. It also summarizes the main concepts of PPA to understand the approach I introduce to measure cross-predictability.

3.1 Long-short Portfolios, Simple Factors and Principal Portfolios

The starting point in KMP is represented by *linear strategies* of the type

$$R^w_{t+1} = w'_t R_{t+1} = \sum_{j=1}^{N} \underbrace{(S'_t L_j)}_{\text{position in } j} \underbrace{R_{j,t+1}}_{\text{return of } j} = S'_t L R_{t+1} \quad (1)$$

$R_{t+1} = (R_{i,t+1})_{i=1}^N \in \mathbb{R}^N$ is a vector of returns in excess of the risk-free rate, $S_t = (S_{i,t})_{i=1}^N \in \mathbb{R}^N$ is a vector of signals (e.g. firm characteristics) and $w_t = (w_{i,t})_{i=1}^N \in \mathbb{R}^N$ is a vector of portfolio weights. $L \in \mathbb{R}^{N\times N}$ is the *position matrix* where each column translates signals into a portfolio position in each assets, $w'_t = S'_t L$. L allows the position $S'_t L_j$ in any

[1]For a comprehensive review of the methods employed in this area, see Giglio et al. (2022). Bagnara (2022) offers a thorough review of the empirical results.

asset j to depend on the signal of *all* assets. This is a general framework nesting several cases based on the specification of L. By controlling L I control in fact the effect of cross-prediction effects in trading strategies.

Consider setting $L = I_N$ where I_N is an $N \times N$ identity matrix. Then, Eq.(1) delivers a "simple factor" $\tilde{F}$ which considers only own-signal predictions:

$$\tilde{F}_{t+1} = \sum_{j=1}^{N} S_{j,t} R_{j,t+1} = S_t' R_{t+1} \tag{2}$$

When the entries off the main diagonal of L are set to zero, I disregard any cross-predictability, and if the main diagonal elements all equal one, the corresponding strategy trades each asset in the same proportion as the strength of its own signal.

Now consider further setting $S_{j,t} = D_{j,t}$ where

$$D_{j,t} = \begin{cases} +1, & \text{if } S_{j,t} = max\{S_{1,t}, S_{2,t}, ..., S_{N,t}\} \\ -1, & \text{if } S_{j,t} = min\{S_{1,t}, S_{2,t}, ..., S_{N,t}\} \\ 0, & \text{if else} \end{cases} \tag{3}$$

A strategy based on the signal $D_{j,t}$ buys the highest-signal asset and sells the one with the lowest signal strength, disregarding the remaining ones. If assets are characteristic-sorted portfolios, I obtain a market-neutral strategy mimicking the risk associated with the signal, i.e., the traditional long-short portfolio LS used in empirical asset pricing:

$$LS_{t+1} = \sum_{j=1}^{N} D_{j,t} R_{j,t+1} = D_t' R_{t+1} \tag{4}$$

Instead, PPs allow the signal of each asset to predict the return of all assets (or to "cross-predict" other assets). The k-th PP can be represented as

$$PP_{t+1}^k = S_t' v_k u_k' R_{t+1} = S_t^{v_k'} R_t^{u_k} = \sum_{j=1}^{N} S_{j,t}^{v_k} R_{j,t+1}^{u_k} \tag{5}$$

where u_k and v_k are, respectively, the k-th column of the orthonormal matrices U and V from the singular value decomposition of the so-called *prediction matrix* $\Pi \in \mathbb{R}^{N \times N}$:

$$\Pi_t := E_t[R_{t+1} S_t'] = U \bar{\Lambda} V' \tag{6}$$

Here, $\bar{\Lambda} = \text{diag}(\bar{\lambda}_1, ..., \bar{\lambda}_N)$ is the diagonal matrix of singular values. The prediction matrix is a very interesting mathematical object. The entries on the main diagonal $\Pi_{i,i}$ represent the expected return of a strategy that chooses an asset's position equal to its own signal strength. Elements off the main diagonal $\Pi_{i,j},\ i \neq j$, capture the return of a strategy that determines asset i's position based on asset j's signal. Hence, Π contains the information relative to the entire asset universe, including cross-prediction terms, which means it is the source of cross-predictability. I address the properties of Π more in detail in Section 4.1.

3.2 An overview of Principal Portfolio Analysis

Before illustrating the methodology used in this paper, I outline the main steps in the derivation of PPs, including their decomposition into positive- and zero-factor exposures, in order to provide the intuition necessary to follow the rest of the paper. The original notation is used where possible. The reader familiar with PPA can skip this section.

KMP maximize the expected return of a linear strategy subject to a constraint on L:

$$\max_{L:||L||\leq 1} E[S_t' L R_{t+1}] \tag{7}$$

The matrix norm is defined as $||L|| = sup\{||Lx|| : x \in \mathbb{R}^m \text{ with } ||x|| = 1\}$ where $||x|| = (\sum_i x_i)^2$ is the standard Euclidean norm of a vector in $x \in \mathbb{R}^N$. The constraint represents a bound on the portfolio size $||L'S_t||$ corresponding to portfolio weights $S_t'L$ that admits only linear strategies with a position size not exceeding the position size of the simple factor $\tilde{F}$, as explained in Appendix A.1. KMP show that the solution to (7) is

$$L = (\Pi'\Pi)^{-1/2}\Pi' = V\bar{\Lambda}^{-1}V'V\bar{\Lambda}U' = \sum_{k=1}^{N} v_k u_k' \tag{8}$$

where I used the singular value decomposition on Π as in Eq.(6). The optimal solution can be decomposed into a collection of linear strategies denoted as PPs:

$$PP_{t+1}^k = S_t' \underbrace{v_k u_k'}_{L_k} R_{t+1} \tag{9}$$

for $k = 1, ..., N$. This strategy trades the portfolio $u_k' R_{t+1}$ based on the signal coming from the portfolio $S_t' v_k$. In contrast to Principal Component Analysis (PCA), which decomposes the variance, PPA decomposes expected returns, which equal the eigenvalue of each PP (see Proposition 4 in KMP)

$$E[PP_{t+1}^k] = tr(\Pi v_k u_k') = tr(U\bar{\Lambda}V' v_k u_k') = \bar{\lambda}_k \tag{10}$$

where tr denotes the trace of a matrix.

PPs can be further decomposed into two sets of portfolios with different factor exposures, namely Principal Exposure Portfolios (PEPs) and Principal Alpha Portfolios (PAPs). First, recall that any matrix $\Pi \in \mathbb{R}^{N\times N}$ has a symmetric part $\Pi^s = \frac{1}{2}(\Pi + \Pi')$ and an antisymmetric part $\Pi^a = \frac{1}{2}(\Pi - \Pi')$ such that $\Pi = \Pi^s + \Pi^a$. Similarly, I can see any linear strategy L as the sum of a symmetric and antisymmetric part, $L = L^s + L^a$. Hence

$$R^w_{t+1} = S'_t L R_{t+1} = S'_t L^s R_{t+1} + S'_t L^a R_{t+1} \tag{11}$$

While the factor risk of a linear strategy is entirely due to its symmetric part, such that the antisymmetric strategy is always factor neutral, both components contribute to its expected return, i.e., the antisymmetric strategy can give rise to alphas, as shown in Appendix A.2. KMP argue that symmetric strategies can be represented more generally as strategies trading assets based on their own signals, similarly to the simple factor $\tilde{F}$. The optimal symmetric strategy is $L^s = (\Pi^s \Pi^s)^{-1/2}\Pi^s$, which can be decomposed into N PEPs:

$$PEP^k_{t+1} = \underbrace{S'_t w^s_k}_{S^{w^s_k}_t} \underbrace{(w^s_k)' R_{t+1}}_{R^{w^s_k}_{t+1}} \tag{12}$$

for $k = 1, ..., N$ where w^s_k is the k-th eigenvector corresponding to the k-th eigenvalue of Π^s. The expected return of the k-th PEP equals the corresponding eigenvalue λ^s_k of Π^s, $E[PEP^k_{t+1}] = \lambda^s_k$ (see Proposition 6 in KMP).

The optimal antisymmetric strategy can be decomposed into $N^a \leq N/2$ PAPs:

$$PAP^k_{t+1} = \underbrace{S'_t x_k (y_k)' R_{t+1}}_{S^{x_k}_t R^{y_k}_{t+1}} - \underbrace{S'_t y_k (x_k)' R_{t+1}}_{S^{y_k}_t R^{x_k}_{t+1}} \tag{13}$$

for $k = 1, ..., N^a$ where x_k and y_k are the real and the imaginary components of the eigenvectors associated with the eigenvalues of $(\Pi^a)'$.[2] Differently from symmetric strategies, which exploit only own-asset signals, PAPs represent long-short strategies trading portfolios $R^{y_k}_{t+1}$ and $R^{x_k}_{t+1}$ against each other based on each other's signal (hence using $S^{x_k}_t$ and $S^{y_k}_t$, respectively), with no factor exposure. The expected return for the k-th PAP equals twice the corresponding eigenvalue λ^a_k of the antisymmetric matrix Π^a, $E[PAP^k_{t+1}] = 2\lambda^a_k$.

To sum up, PPs are optimal linear strategies that simultaneously account for own-asset predictions and cross-asset predictions, which are captured by PEPs and PAPs, respectively.

4 Methodology

After giving a quick overview of PPA, I now explain the approach I employ to build a performance-based measure of cross-predictability that allows to decompose it into cross-predictive signals and cross-predicted returns.

4.1 A Closer Look at the Prediction Matrix

Let us start by examining the prediction matrix Π_t for a given firm characteristic S^k_t:

$$\Pi_t = E_t[R_{t+1}S_t'] = E_t \begin{bmatrix} R_1S_1 & R_1S_2 & \cdots & R_1S_N \\ R_2S_1 & R_2S_2 & \cdots & R_2S_N \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 & R_{N-1}S_2 & \cdots & R_{N-1}S_N \\ R_NS_1 & R_NS_2 & \cdots & R_NS_N \end{bmatrix}$$

where I omit the time index for matrix entries for ease of exposition. The corresponding symmetric and antisymmetric matrices Π^s_t and Π^a_t are

[2] See Appendix A.3 for further details.

$$\Pi_t^s = \frac{1}{2}(\Pi_t + \Pi_t') = \frac{1}{2}E_t \begin{bmatrix} 2R_1S_1 & R_1S_2 + R_2S_1 & \cdots & R_1S_N + R_NS_1 \\ R_2S_1 + R_1S_2 & 2R_2S_2 & \cdots & R_2S_N + R_NS_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 + R_1S_{N-1} & R_{N-1}S_2 + R_2S_{N-1} & \cdots & R_{N-1}S_N + R_NS_{N-1} \\ R_NS_1 + R_1S_N & R_NS_2 + R_2S_N & \cdots & 2R_NS_N \end{bmatrix}$$

$$\Pi_t^a = \frac{1}{2}(\Pi_t - \Pi_t') = \frac{1}{2}E_t \begin{bmatrix} 0 & R_1S_2 - R_2S_1 & \cdots & R_1S_N - R_NS_1 \\ R_2S_1 - R_1S_2 & 0 & \cdots & R_2S_N - R_NS_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 - R_1S_{N-1} & R_{N-1}S_2 - R_2S_{N-1} & \cdots & R_{N-1}S_N - R_NS_{N-1} \\ R_NS_1 - R_1S_N & R_NS_2 - R_2S_N & \cdots & 0 \end{bmatrix}$$

The entries along the main diagonal of Π_t contain only own-asset prediction terms. With Π_t as a diagonal matrix, in fact, I are back in the case of the simple factor $\tilde{F}_t$ (see Section 3.1). Entries off the main diagonal, instead, contain cross-prediction terms, as the signal of one asset at time t is paired with the return of another asset at time $t+1$. Based on this observation, I can disentangle the economic value of cross-prediction effects by manipulating the prediction matrix. By performing three different transformations, I can capture three types of cross-prediction effects, as explained in the following.

4.2 Cross-predictive Signals and Cross-predicted Returns

What kind of cross-prediction effects can I expect? Consider a simple example illustrated in Figure 1 with $N = 3$ stocks and a single signal $S_t = (S_{1,t}, S_{2,t}, S_{3,t})$ used to predict the returns $R_{t+1} = (R_{1,t+1}, R_{2,t+1}, R_{3,t+1})$. Each signal $S_{j,t}$ predicts the return of the same stock

$R_{j,t+1}$, but it can also predict the returns of the other two stocks. Hence, there are three arrows originating from each $S_{j,t}$, one for each return $R_{j,t+1}$, for $j = 1, ..., N = 3$. Horizontal lines represent own-asset predictions. Focusing on the remaining ones, I can represent cross-prediction effects differently depending on the point of view I take. If I focus on signals (left side of the chart), I can think of cross-predictability as the predictive power that each signal $S_{j,t}$ has to predict other assets' returns, $R_{i,t+1}$ with $i \neq j$. I name this *cross-predictive signals*. If, instead, I focus on returns (right side), cross-predictability is represented by the amount of one asset's return $R_{j,t+1}$ captured by the signals of other assets, $S_{i,t}$ with $i \neq j$. I name this *cross-predicted returns*. Notice that considering all cross-predictive signals for $j = 1, ..., N$ is the same as considering all cross-predicted returns for $j = 1, ..., N$, as both capture any possible cross-prediction effect. The union of the cross-predictive signals and the cross-predicted returns of asset j captures any cross-effects involving that asset, which means either originating from its own signals or from other assets. I name the whole effect simply *cross-prediction total*. The special structure of the prediction matrix allows us to isolate these three effects precisely and to quantify their economic value, as shown below.

4.2.1 Cross-predictive Signals

PPA uses one signal at a time. For a given signal S, consider excluding the N-th cross-predictive signal, which is the effect that firm $j = N$'s characteristic S_N has on other firms' returns, without restricting its own-predictability effect. Through the lens of the prediction matrix Π_t, this means setting $S_{t,N} = 0$, excluding the entry on the main diagonal.[3] Hence:

$$\dot{\Pi}_t = E_t \begin{bmatrix} R_1 S_1 & R_1 S_2 & \cdots & 0 \\ R_2 S_1 & R_2 S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1} S_1 & R_{N-1} S_2 & \cdots & 0 \\ R_N S_1 & R_N S_2 & \cdots & R_N S_N \end{bmatrix}$$

$$\dot{\Pi}_t^s = \frac{1}{2} E_t \begin{bmatrix} 2R_1 S_1 & R_1 S_2 + R_2 S_1 & \cdots & R_N S_1 \\ R_2 S_1 + R_1 S_2 & 2R_2 S_2 & \cdots & R_N S_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1} S_1 + R_1 S_{N-1} & R_{N-1} S_2 + R_2 S_{N-1} & \cdots & R_N S_{N-1} \\ R_N S_1 & R_N S_2 & \cdots & 2R_N S_N \end{bmatrix}$$

$$\dot{\Pi}_t^a = \frac{1}{2} E_t \begin{bmatrix} 0 & R_1 S_2 - R_2 S_1 & \cdots & -R_N S_1 \\ R_2 S_1 - R_1 S_2 & 0 & \cdots & -R_N S_2 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1} S_1 - R_1 S_{N-1} & R_{N-1} S_2 - R_2 S_{N-1} & \cdots & -R_N S_{N-1} \\ R_N S_1 & R_N S_2 & \cdots & 0 \end{bmatrix}$$

The notation $\dot{\Pi}_t$ indicates that Π_t has been manipulated to exclude cross-predictive signals involving one asset. This transformation does not alter the matrix properties that are relevant for PPA. In particular, $\dot{\Pi}_t = \dot{\Pi}_t^s + \dot{\Pi}_t^a$, and $\dot{\Pi}_t^s$ and $\dot{\Pi}_t$ are still symmetric and antisymmetric, respectively, and all $\dot{\Pi}_t$, $\dot{\Pi}_t^s$, $\dot{\Pi}_t^a$ are of full rank and thus invertible (unless $R_N = 0$, which is usually not the case, which would make even the original Π_t rank-deficient). This implies that the number of non-zero eigenvalues of all $\dot{\Pi}_t$, $\dot{\Pi}_t^s$, $\dot{\Pi}_t^a$ does not change with respect to the baseline case, and all the results in Kelly et al. (2023) are still valid. The

[3]The procedure of setting a column (or row as below) to evaluate the impact of a certain covariate on a PCA-based methodology is used also in Kelly et al. (2019) to test instrument significance.

crucial step is not to set main diagonal elements to zero. Mathematically, this would create rank-deficient matrices. Economically, this would capture own-asset predictability effects, which is not what I are after.

4.2.2 Cross-predicted Returns

Now consider excluding the N-th cross-predicted return, which is the effect that all other firms' characteristics S_i, $i = 1, ..., N,\ i \neq j$ have on firm $j = N$'s return, without restricting its own-predictability effect. Working on Π_t, this means setting $R_{N,t} = 0$ outside the main diagonal. Hence:

$$\ddot{\Pi}_t = E_t \begin{bmatrix} R_1S_1 & R_1S_2 & \cdots & R_1S_N \\ R_2S_1 & R_2S_2 & \cdots & R_2S_N \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 & R_{N-1}S_2 & \cdots & R_{N-1}S_N \\ 0 & 0 & \cdots & R_NS_N \end{bmatrix}$$

$$\ddot{\Pi}_t^s = \frac{1}{2} E_t \begin{bmatrix} 2R_1S_1 & R_1S_2 + R_2S_1 & \cdots & R_1S_N \\ R_2S_1 + R_1S_2 & 2R_2S_2 & \cdots & R_2S_N \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 + R_1S_{N-1} & R_{N-1}S_2 + R_2S_{N-1} & \cdots & R_{N-1}S_N \\ R_1S_N & R_2S_N & \cdots & 2R_NS_N \end{bmatrix}$$

$$\ddot{\Pi}_t^a = \frac{1}{2} E_t \begin{bmatrix} 0 & R_1S_2 - R_2S_1 & \cdots & R_1S_N \\ R_2S_1 - R_1S_2 & 0 & \cdots & R_2S_N \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 - R_1S_{N-1} & R_{N-1}S_2 - R_2S_{N-1} & \cdots & R_{N-1}S_N \\ -R_1S_N & -R_2S_N & \cdots & 0 \end{bmatrix}$$

The notation $\ddot{\Pi}_t$ indicates that Π_t has been manipulated to exclude cross-predicted returns involving one asset. Analogously to the first transformation, this does not alter the properties of the prediction matrix that are relevant for PPA.

4.2.3 Cross-prediction Total

Finally, consider excluding the N-th cross-prediction total, which amounts to disregard any type of cross-predictability linked to firm N, again without restricting its own-predictability effect. From the point of view of the prediction matrix Π_t, this means setting $S_{t,N} = 0$ and $R_{N,t} = 0$ excluding the entry on the main diagonal. In this case:

$$\dddot{\Pi}_t = E_t \begin{bmatrix} R_1S_1 & R_1S_2 & \cdots & 0 \\ R_2S_1 & R_2S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 & R_{N-1}S_2 & \cdots & R_{N-1}S_N \\ 0 & 0 & \cdots & R_NS_N \end{bmatrix}$$

$$\dddot{\Pi}_t^s = \frac{1}{2} E_t \begin{bmatrix} 2R_1S_1 & R_1S_2 + R_2S_1 & \cdots & 0 \\ R_2S_1 + R_1S_2 & 2R_2S_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 + R_1S_{N-1} & R_{N-1}S_2 + R_2S_{N-1} & \cdots & 0 \\ 0 & 0 & \cdots & 2R_NS_N \end{bmatrix}$$

$$\dddot{\Pi}_t^a = \frac{1}{2} E_t \begin{bmatrix} 0 & R_1S_2 - R_2S_1 & \cdots & 0 \\ R_2S_1 - R_1S_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ R_{N-1}S_1 - R_1S_{N-1} & R_{N-1}S_2 - R_2S_{N-1} & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

The notation $\dddot{\Pi}_t$ indicates that Π_t has been manipulated to exclude cross-predictive signals and cross-predicted returns involving one asset. As with the previous two transformations, the matrix properties of interest for PPA are preserved.[4]

4.3 The Economic Value of Cross-Predictability

Equipped with this straightforward way to distinguish different types of cross-prediction effects, I now turn to how to quantify their economic value based on expected returns and SRs, two classical performance measures that fit well the PPA optimization framework.

Starting from expected returns, for any asset $j = 1, ..., N$, I perform one of the transformations explained above to capture cross-predictive signals, cross-predicted returns or cross-prediction total. Then, I compute the average realized out-of-sample (OOS) return of the first PP in each case and I calculate the negative percentage deviation from the baseline case without restrictions on Π.[5] Three cross-prediction measures are obtained:

$$\text{value of Cross-predictive Signal } j := CPS_j^{\bar{R}} = 1 - \frac{\bar{R}_{-column_j}}{\bar{R}_{full}} \tag{14}$$

$$\text{value of Cross-predicted Return } j := CPR_j^{\bar{R}} = 1 - \frac{\bar{R}_{-row_j}}{\bar{R}_{full}} \tag{15}$$

$$\text{value of Cross-prediction Total } j := CPT_j^{\bar{R}} = 1 - \frac{\bar{R}_{-column_j,-row_j}}{\bar{R}_{full}} \tag{16}$$

[4]Notice that by applying any of the three transformations for all firms $j = 1, 2, ..., N$ collapses the prediction matrix Π_t to a diagonal matrix populated with only own-predictability terms, i.e., I are back in the case of the simple factor $\tilde{F}$.

[5]A strategy using only the first PP is not necessarily the optimal one. Any portfolio with positive eigenvalue should be traded. With PEPs, the optimal strategy shorts negative-eigenvalue portfolios, if any. With this in mind, I focus on the first PP for two reasons. First, this portfolio has the highest expected return among all PPs, and empirically the subsequent ones have singular values and average returns close to zero. Hence, trading only the first PP does not lead to big losses in performance. Second, this is beneficial for interpretation purposes. The LS factor, the standard in the literature, and the simple factor $\tilde{F}$, are both single portfolios. To remain close in spirit to them, I exclude strategies composed of multiple PPs.

where $\bar{R}_{full}$ denotes the average return of the baseline case and $\bar{R}_{-column_j}$, $\bar{R}_{-row_j}$ and $\bar{R}_{-column_j,-row_j}$ denote the average return after zeroing out the j-th column, row, and column and row, respectively, of the prediction matrix Π, except for the j-th element of the main diagonal. In sequential order $CPS_j^{\bar{R}}$, $CPR_j^{\bar{R}}$ and $CPT_j^{\bar{R}}$ measure the percentage loss in expected return from the first PP resulting from disregarding cross-prediction effects from, to, and both from and to asset j. I repeat the same procedure for the average return of the first PEP and PAP to obtain analogous measures.

From Eq.(10), I know that the k-th singular value of Π is also the expected return of the corresponding k-th PP. As an additional estimate of the cross-predictability value based on expected returns, in Appendix B I compute percentage deviations between singular values before and after performing the transformations introduced above.

I also provide measures based on annualized OOS SRs. Accounting for portfolio risk as represented by the standard deviation in addition to expected returns delivers a unit-free measure for the economic value of cross-predictability. The SR-based measures are

$$\text{value of Cross-predictive Signal } j := CPS_j^{SR} = 1 - \frac{SR_{-column_j}}{SR_{full}} \tag{17}$$

$$\text{value of Cross-predicted Return } j := CPR_j^{SR} = 1 - \frac{SR_{-row_j}}{SR_{full}} \tag{18}$$

$$\text{value of Cross-prediction Total } j := CPT_j^{SR} = 1 - \frac{SR_{-column_j,-row_j}}{SR_{full}} \tag{19}$$

where the interpretation of the pedices is analogous to one given above. These measures capture the percentage loss in the annualized OOS SR resulting from excluding cross-prediction effects among the base assets used to build PP strategies.[6]

[6]Calculating percentage differences among SRs is a common practice in the literature, see e.g. Campbell

Regardless of whether they focus on expected returns or on SRs, CPS, CPR and CPT are performance-based measures that capture the economic impact of ignoring cross-prediction effects when implementing a trading strategy. Compared to previous studies, which mainly build long-short portfolios based on the lead-lag relations, my measures have the advantage of identifying precisely whether it is more important to predict a certain part of the cross-section with other assets (CPR) or if cross-prediction effects are mainly due to the signal from a part of the cross-section that is very informative for other assets (CPS). Remaining in the framework of PPA, a very recent and promising methodology, I are able to provide clean-cut measures of cross-predictability that do not suffer from confounding variables affecting the simple outcome of long-short portfolios, as I can precisely restrict only the cross-prediction terms of the prediction matrix, leaving the rest unchanged. Benchmarking the performance of PPs that exclude certain cross-effects to the "full" strategy, my approach keeps all the rest fixed and thus does not need to control for further omitted variables.

5 Data

I apply my approach to the data from Jensen et al. (2022). The database contains 153 signals from the asset pricing literature at the individual stock-month level.[7] The period considered is January 1968 - December 2019. Characteristic-sorted portfolios are formed as detailed below thoroughly following KMP. When forming portfolios (each month), nano and micro stocks are excluded, and the minimum number of stocks per portfolio is restricted to 10. Portfolio returns and portfolio-level signals are computed as weighted averages of

and Thompson (2008).

[7] I thank Bryan Kelly for sharing the data. Appendix C reports a list of the signals used.

the portfolio components, using capped-value weights at the 80$^{\text{th}}$ percentile of the NYSE breakpoints.[8] The prediction matrix is estimated via 120-month-window rolling averages of $R_{t+1}S_t'$, as in KMP. PPA is applied with the same frequency for one-month ahead OOS return predictions.

6 Empirical Results

In this section I apply PPA to size-sorted portfolios and describe the empirical application of the cross-predictability value measures, focusing on their economic interpretation.

6.1 Size-sorted Portfolios: Size as signal

Among the many paradigms that can be used to group individual stocks into portfolios, characteristic-sorted portfolios are the most frequently used. For instance, the portfolios made available by Kenneth French have been used for decades as test assets for many models, becoming universal benchmarks in asset pricing.[9] Here, I sort stocks into ten groups based on size. Previous research has documented lead-lag effect from large to small stocks (Lo and MacKinlay, 1990). Moreover, small stocks are often problematic. Hou et al. (2020) find that 65% of anomalies are insignificant after excluding microcaps, and modern methods perform remarkably different when applied to large or small stocks (Gu et al., 2020). In other words,

[8] Breakpoints are provided at https://www.bryankellyacademic.org/. Portfolio-level characteristics are cross-sectionally ranked and mapped into the $[-0.5, 0.5]$ interval before applying PPA, as it is standard in ML applications (e.g. Kelly et al. (2019)). Moreover, as in Section V.A in KMP, returns are cross-sectionally demeaned to focus predictions on cross-section differences in returns rather than on time-series fluctuations in their common market component. In words, this is equivalent to trading assets with long positions hedged by shorting an equal-weighted average of all portfolios. KMP show that results are very similar if the market return is not hedged in this way, and thus I follow their main approach here. Notice that this procedure makes it equivalent to use either returns or returns in excess of the risk-free rate.

[9] See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

firm market capitalization gives rise to interesting patterns that are worth investigating, hence the choice of size-sorted portfolios. Considering the full ten-portfolio partition allows to identify potential cross-prediction effects involving not only the characteristic-based extremes of the cross-section, but also those regarding the middle range of the distribution, which is typically disregarded in the literature without an explicit explanation.

Using these ten size-sorted portfolios, I compute the long-short portfolios LS, which is a "classical" size factor, the simple factor $\tilde{F}$ and the first PP, PEP and PAP. As mentioned above, I focus only on the first portfolio for PPA and disregard strategies composed of multiple portfolios. The comparison among these different assets have meaningful economic interpretations when considering Section 3.1. Differences between the performance of LS and $\tilde{F}$ reveal the economic value contained in the signals in the middle of the cross-section instead of focusing only on the extremes. Comparing the first PP and $\tilde{F}$, instead, captures the additional value of cross-prediction effects beyond own-asset predictions. Since PPA must be performed one signal at a time, I start by using size to illustrate the main logic of the analysis. In Section 6.2, I extend the approach to all signals available in the dataset.

6.1.1 Singular Values and Average Returns

Similarly to KMP, let us start by looking at how PPA performs for size-sorted portfolios. Figure 2 shows time-averages of the singular values of Π, Π^s and Π^a on the left and the average OOS returns of the corresponding PPs, PEPs and PAPs on the right, with ± 2 standard errors confidence bands.[10] First, I see that overall the singular values and eigenvalues on the left panels are very close to the average returns on the right panels, as I would expect from the

[10] As explained in Section 5, PPA is performed on a rolling window of 120 months, which gives rise to a time-series of singular values for each component of the prediction matrix.

theory discussed in Section 3.2.[11] In Panel D, the average returns are all basically zero except for the first one. Thus, focusing just on the first PP is the optimal choice. In Panel B there are two relevant PEPs, i.e., the first one and the last one, with positive and negative average returns, respectively. Since PEPs represent factor exposure, this phenomenon is expected as I recognize something similar to the two legs of the classical size factor.[12] Panel C shows that among the portfolios with zero factor exposures, only the first one earns a positive expected return. The overall message of Figure 2 is that in the cross-section of size-sorted portfolios there is one important latent factor exploiting cross-prediction effects (PP). Furthermore, there is one alpha-generating portfolio (PAP) and two additional portfolios exposed to the size factor (PEPs) with positive and negative expected return, respectively.

6.1.2 PPA Performance

To evaluate the performance of PPA, I follow KMP and report in Figure 3 the SR of the first PP, the first PEP, and the first PAP. As a benchmark, I also show the SR of the simple factor. The dashed line is the SR of the LS portfolio. Each bar is accompanied by the ± 2 approximate standard error bar based on Lo (2002). The PP strategy outperforms both $\tilde{F}$ and LS, which means that cross-prediction effects contain valuable information for investment purposes. Both PEP and PAP perform better than LS, but PAP earns the highest SR. Put differently, exploiting only cross-predictability but no factor risk reveals the best choice. All PPA-based strategies generate a significant SR while $\tilde{F}$ does not. Notably,

[11] Since I are comparing realized eigenvalues to OOS average returns I cannot expect a perfect match. This is also what happens in KMP.

[12] The optimal symmetric strategy buys all portfolios with positive eigenvalues and sells those with negative eigenvalues, which would result in a long-short portfolio in this case. Here, I focus only on the first PEP for ease of interpretation to get a solid grasp of the sources of cross-predictability.

LS does better than $\tilde{F}$.[13] This means that accounting for the information from some parts of the cross-section (the middle) can be actually detrimental even in the case of own-asset predictions. It could be that the market capitalization of mid-cap portfolios is a poor signal for their own returns. A similar intuition is confirmed later on.

Before moving on to cross-predictability measures, in Table 1 I report the correlation matrix of LS, $\tilde{F}$, PP, PEP and PAP (excluding the last row) to highlight some interesting facts. The first observation is the strikingly high correlation between LS and $\tilde{F}$, which are essentially the same factor up to a sign. However, trading portfolios on signal strength instead of simply focusing on a long-short strategy hinders portfolio performance. Another interesting correlation is the one between PEP and $\tilde{F}$, which is 0.79. This is line with the theory, that shows that PEPs are the only responsible for factor risk in PPA. PAP, instead, shows the lowest correlation with $\tilde{F}$ as it is a zero-exposure portfolio. The last row reports the average R^2 obtained by regressing each size-sorted portfolio on a one-factor model corresponding to the factor in the column. PPs (including PEP and PAP) produce much lower fits compared to the standard LS factor. In other words, a better investment performance is not accompanied by a higher cross-sectional explanatory power. This result is noteworthy for two reasons. First, it is similar to the findings in Kelly et al. (2022), who document that the OOS R^2 is an incomplete measure of the economic value of a model, and that substantial profits can be generated even with a low fit. Second, it suggests that PPA-based strategies might produce "weak" factors (Lettau and Pelger, 2020) that are not

[13] I adjust the sign of LS such that it has a positive mean return to be consistent with what is typically expected from the literature. For size-sorted portfolios, LS means actually buying small stocks and selling large stocks. I do not do this for $\tilde{F}$ because this would introduce a look-ahead bias that is less of an issue for LS factors that have been documented for decades. In any case, even flipping the sign of $\tilde{F}$ would not change the conclusion that it performs worse than LS.

part of the stochastic discount factor but are nevertheless priced as non-diversifiable risk.

6.1.3 Cross-predictability: Average Returns

I now turn to measuring the economic value of cross-predictability. I start by using $CPS_j^{\bar{R}}$, $CPR_j^{\bar{R}}$ and $CPT_j^{\bar{R}}$ for $j = 1, ..., N$. I stress that j here denotes the base asset as opposed to k, which refers to the ordering of PPs. I compute these statistics for the first PP and report them on the top-left panel of Figure 4. One of the most apparent patterns is that there is essentially no value in the center of the cross-section, because all three cross-prediction measures are almost zero. The extremes are instead the most important parts, as informally argued in the literature (e.g. Lettau and Pelger (2020)). This fact is also in line with the idea of using the extremes as the two legs of a long-short portfolio mimicking the underlying risk factors, a widely used practice. At the extremes, the economic value of cross-predictability concentrates on small stocks. Excluding any cross-prediction regarding this portfolio results in a loss of roughly 57% of the average PP return. The same operation involving large stocks leads to a loss of around 12%. Hence, small firms are the most relevant part of the cross-section from a cross-predictability perspective. Where does its economic value come from? Mainly from cross-predicted returns. In fact, predicting small stocks with signals from other assets creates much larger economic value ($CPR_1^{\bar{R}} = 0.28$) than using small companies' signals to predict other assets ($CPS_1^{\bar{R}}$ close to zero).

Interestingly, the opposite holds for large stocks. Roughly 11% of PP average return is lost once I do not consider the information that large firms' signals have for other assets. However, excluding cross-prediction from other assets towards large stocks does not lead any loss ($CPS_{10}^{\bar{R}} = -0.03$). These facts resemble the lead-lag effects documented in Lo and

MacKinlay (1990), where past returns of large stocks systematically lead small stocks returns but not vice-versa. Here, the difference is that I are using size as signal to predict returns instead, and the signal is measured at the same time for both group of stocks.

In the second and third panel on the left of Figure 4 I repeat the same exercise for the first PEP and the first PAP, respectively. The shape of the cross-prediction measures for symmetric and antisymmetric strategies is somewhat different, but the general patterns highlighted above persist: the extremes of the distributions contain most of the economic value of cross-predictability and the value from predicting an asset's return from other assets decreases with size. Notably, these patterns hold to a great extent also for PEP. Although the cross-prediction measures are remarkably high not only for the first size-sorted portfolio but also for the second and the third, I have to keep in mind that PEPs are symmetric strategies that drive the factor exposure of a certain trading strategy exploiting own-asset signals. However, looking at the structure of Π^s in Section 4, I see that cross-prediction terms actually exist and thus eventually impact the matrix eigenvalues and expected returns. Hence, neglecting cross-predictability impacts also PEP strategies. As expected from the theory, the third panel of Figure 4 reflects very closely the first one. PAPs are indeed long-short strategies trading one asset against the other based on the other asset's signal. As such, they exploit in full cross-predictability patterns and are thus likely to be even more sensitive to neglecting cross-prediction effects. The $CPT_1^{\bar{R}}$ is indeed 0.69, even higher than the same measure for the PP case. Cross-predicting large stocks is not detrimental any more, but disregarding this effect leads only to a modest performance loss ($CPR_{10}^{\bar{R}} = 0.07$).

In PPA, the eigenvalues of the prediction matrix (including its symmetric and antisymmetric components) correspond to the expected returns of PPs. However, as explained in

KMP, eigenvalues will never perfectly match realized average returns out-of-sample but will follow them very closely, as can be seen from Figure 3. Abstracting from few small discrepancies, in Appendix B I observe a strong similarity also between cross-predictability measures based on average returns and on eigenvalues. As eigenvalues are only the theoretical counterpart of expected returns, and the aim of this study is to provide an actual performance-based measure of cross-predictability, for the remainder of the paper I will focus on average-return-based measures.

6.1.4 Cross-predictability: Sharpe Ratios

The right side of Figure 4 illustrates CPS_j^{SR}, CPR_j^{SR} and CPT_j^{SR} for $j = 1, ..., N$. Results are shown for the first PP, PEP and PAP, from top to bottom, respectively. Similarly to what is observed for average returns, small stocks are the most important part of the cross-section: excluding any type of cross-prediction effects involving the first size-sorted portfolio results in a loss of almost 50% of the first PP SR. Most of this effect is captured by the information contained in other assets predicting small stocks ($CPR_1^{SR} = 0.24$), rather than vice-versa ($CPS_1^{SR} = -0.01$). Using SRs, it must be noticed that CPS is close to zero for all portfolios, which means on average no portion of the cross-section alone has predictive power for the entire remaining part. However, when put together, cross-prediction effects are strong for small companies, but rapidly decrease to zero in the middle of the distribution until becoming negative for the larger half of stocks ($CRV_{10}^{SR} = -0.06$).

These patterns hold generally also for the other two investment strategies. A couple of observations are noteworthy nonetheless. In particular, for PEP, CPR^{SR} is never positive except only for the smallest group of stocks. PAP are remarkably similar to PP and, as before,

CPR is positive for large firms ($CPR^{SR}_{10} = 0.11$), but not nearly close to the magnitude of smallest stocks ($CPR^{SR}_{1} = 0.63$). The comparison across the panels of the figure helps understanding the origin of this finding. The component of returns of large stocks due to exposure to size risk cannot really be explained by other assets' size (middle panel), while a portion of the alpha component can be meaningfully cross-predicted by other firms (bottom panel). The first effect dominates the second one, such that overall trying to cross-predict large stocks does not improve portfolio performance or even hinders it (top panel).

To sum up, cross-prediction measures tell us that it is crucial not to ignore cross-predictability for investment purposes, but only for smaller stocks, while larger stocks appear too hard to predict using the other assets.

6.2 Size-sorted Portfolios: All signals

PPA can be performed only one signal at a time. After illustrating the results using size as a signal, I extend the analysis to all 153 signals in the dataset for the same ten size-sorted portfolios and average out the results to investigate the robustness of out findings.

Figure 5 shows the average SR of the first PP, PEP and PAP over all signals, together with the size LS portfolio (independent of the signal) and the average simple factor $\tilde{F}$. Results are very close to those observed above: each PPA-based strategy outperforms both LS and $\tilde{F}$, and the best outcomes are achieved with PAP. The first PEP now does slightly better than PP, instead of worse as for the single size signal. Investing in each asset based on own-asset signal ($\tilde{F}$) is still strictly worse than a standard long-short strategy (LS).

Figure 6 reports on the left $CPS^{\bar{R}}_j$, $CPR^{\bar{R}}_j$ and $CPT^{\bar{R}}_j$ for $j = 1, ..., N$, averaged over the

153 signals, for PP, PEP and PAP from top to bottom, respectively. As in the single size case, most of the cross-predictive information concentrates on the extremes, and small stocks are the most relevant. The average return of the first PP drops by almost 60% if one excludes any cross-prediction effect involving the first size-portfolio. This effect is dominated by CPR rather than by CPS, while the opposite holds for larger stocks (ninth and tenth portfolio). The contribution of cross-predictability for the rest of the cross-section is negligible. The results relative to PEP and PAP are also remarkably close to what observed in Figure 4.

The right side of the same figure illustrates analogous averages based on CPS_j^{SR}, CPR_j^{SR} and CTP_j^{SR} for $j = 1, ..., N$. The disproportionate importance of small stocks is apparent once again: on average, excluding cross-prediction effects involving the first size-portfolio leads to a decrease in SR of more than 60% for the first PP and around 40% for the first PAP, and more than 20% for PEP. The general decrease in economic value of cross-predictability with market capitalization is present in all three cases, and overall $CPR_1^{SR} > CPS_1^{SR}$ but $CPR_{10}^{SR} < CPS_{10}^{SR}$. All cross-predictability measures are negative for the largest-stock portfolio for symmetric strategies exposed to characteristic factors (PEPs), while their cross-predictive power is positive for factor-neutral antisymmetric strategies (PAP).

6.2.1 Cross-predictability: Portfolio Composition

To shed further lights on the previous findings, now I analyze the optimal portfolio composition. By looking at how the weights assigned by PPA to size portfolios change when restrictions regarding cross-prediction effects are imposed on Π compared to the unrestricted case, I can get a better idea about the determinants of cross-predictability. Figure 7 shows time-series averages of the portfolio weights composing the first PP, PEP and PAP, from left to right, averaged over 153 signals. For PP, they take the form $S_t' v_1 u_1'$.[14] The bars represent the weights assigned to each size-sorted portfolio after excluding cross-predictive signals, cross-predicted returns or cross-prediction total relative to the smallest (top panels) or largest companies (bottom panels) with the manipulations on Π introduced above.

I focus on the extremes of the distribution because I have seen that they are the most interesting ones. The dashed line represents the portfolio weights with no restrictions ("full Π" case) and coincides in panels belonging to the same PPA strategy. The first PP loads strongly on small stocks, with a weight of 0.35, and essentially shorts the rest of the cross-section, with the exception of portfolios 7, 8 and 9, which have, however, almost no importance. The largest negative weights are for portfolios 2 and 3 (-0.09 and -0.12, respectively). The weight for largest firms is instead -0.05. Overall, the first PP is hardly composed of any "middle-size" stocks, which once again dovetails with the center of the distribution having negligible importance for prediction purposes. An interesting comparison involves LS and $\tilde{F}$. PP assigns a much larger weight to small stocks than to large ones compared with LS, which is equally divided between long and short leg. Such disproportionate weight distinguishes PP also from $\tilde{F}$, which weighs each asset according to the strength of its signal, and reveals decisive based on the performance ratios shown above.

How does the first PP composition change when restricting cross-predictions regarding the first size-portfolio? Excluding small stocks' cross-predictive signals (blue bars) does not change much their weight. The weights for the rest of the cross-section barely change (less than 0.01 in absolute value on average). The result is not surprising: as I have seen, CPS_1 is

[14] Following KMP, weights vary between [-1,1] and sum up to zero in each month. However, taking their average over time, which is necessary for expository purposes, does not ensure the latter property anymore.

generally low, and since PPs are the linear strategies that maximize expected returns, their composition is not strongly impacted by eliminating cross-prediction patterns from small stocks to other stocks' returns. The situation changes more visibly when I do not allow small stocks' returns to be predicted by other assets' signals (orange bars). The weight given to the first portfolio drops to less than half at 0.15. Since the cross-predictive power of small stocks is small, excluding their cross-predicted returns leaves room for them to contribute to PPs almost only through own-asset predictions. The result of this change is a great drop in performance, as shown by the CPR in the previous pictures. Eliminating small stocks' cross-prediction total exacerbates this effect even further, with their weight plummeting to 0.05. Nonetheless, this remains still the third highest weight in absolute value after the second and third portfolio: the simple possibility of predicting small stocks' return with their own signal contributes still strongly to the overall PP performance, a result that I confirm also below. This hints at the fact that low-cap companies have a great room for predictability compared to others, such as including even a one-way prediction mechanism (their own signal) is beneficial for investment outcomes. These findings are in line with with the fact that small stocks are notoriously difficult to arbitrage.

Let us now focus on the bottom-left panel. Based on Section 6.2, I expect different phenomena once I impose restrictions related to the tenth portfolio. Indeed, disregarding large stocks' cross-predictive signals shrinks the overall weights in absolute value, but not substantially: $CPS_{10}^{\bar{R}}$ and CPS_{10}^{SR} are not very high, the cross-predictive power of large companies contributes to portfolio performance only moderately. Excluding cross-prediction effects from other assets' signals for large stocks' returns also does not lead to noticeable differences, except for large firms themselves, whose weight drops almost to zero. In fact,

$CPR^{\bar{R}_{10}}$ and CPR_{10}^{SR} are low or even negative. The same happens when total cross-prediction effects are removed: if large stocks are allowed only to predict themselves, PPA does not identify them as meaningful contributors to expected returns, because the portion of their returns that can be predicted, which is likely low as large stocks are typically liquid and more prone to react promptly to news, might already be captured by other assets' signals.

What happens to the alpha and beta components of PPs? Let us start from PEPs, represented in the two central panels in Figure 7. In the unrestricted case, the first PEP loads positively on larger stocks (last five portfolios) and negatively on smaller stocks, with the exception of the first size-portfolio where the weight is 0.12. Disregarding the smallest 10% of stocks, the first PEP is therefore very similar to the simple portfolio $\tilde{F}$ (correlation of 0.79 in Table 1).[15] When ruling out cross-predictive effects from small stocks to other assets, the weights of all other portfolios decrease in absolute value, while they increase for the first one: after eliminating cross-predictive signals from small stocks, which are indeed detrimental for PEP strategies as from Figure 6, these assets can only predict themselves or being predicted by others, and since $CPR_1^{\bar{R}}$ and CPR_1^{SR} are strongly positive, PEP allocates them a higher weight to improve expected returns. When excluding small companies' cross-predicted returns (orange bars) or cross-prediction totals (black bars), the weight of portfolio 1 drops considerably and PPA assigns an even more negative weight to smaller stocks up to portfolio 5 but also reduces the weight of larger firms from portfolio 6 to 10. When I do not allow large stocks' size to predict other assets' returns, their weight becomes negative and the rest of the weights shrink, indicating that large firms cross-predict a portion of other

[15] I stress that $\tilde{F}$ has not been adjusted in sign to make its average return positive, hence the sign of the correlation with PEP.

stocks' returns even with PEPs. Preventing large stocks from being predicted by others has barely any effect of the rest, hence the total effect is largely dominated by the blue bars.

Finally, let us consider the first PAP in Figure 7 in the last panel. Overall, PAP resembles the simple factor $\tilde{F}$ with the addition of another portfolio that heavily buys small stocks. Excluding the cross-predictive signals of small companies does not change the general composition considerably. It is much more interesting to look at the orange bars instead: if small firms cannot be cross-predicted by the rest of the cross-section, their importance drops almost to zero. Considering that PAPs are strategies that do not contain any own-asset prediction effects, this finding confirms once again that small stocks have close to zero cross-predictive power but it is beneficial to cross-predict them, which fits together with the idea of small stocks containing higher degree of predictable variation due, for example, to trading frictions, a fact widely accepted in the literature. In line with this interpretation, the weight given to larger companies changes dramatically, flipping sign for portfolios 6 to 9 and reaching -0.08 for the last portfolio, which becomes the second most important: the top 10% largest stocks contribute significantly to the portfolio performance of strategies based exclusively on cross-prediction effects, and strikingly so through their effect on small stocks. The final result is essentially a portfolio shorting (with different intensities) the entire cross-section, i.e., a sort of negative market factor. Similar results arise when restricting both types of cross-prediction effects (black bars).

The findings of this section seem to point at an important fact: large stocks contribute to PPs by cross-predicting other assets, especially small stocks, although their predictive power is quite low, such that neglecting it does not change substantially the PPs' composition. Small stocks' contribution, the highest in the entire cross-section, is instead due to themselves being predicted by other firms or simply predicting themselves, probably because they are assets that have more room for return predictability. I confirm these results in Section 7.

6.2.2 Cross-predictability: Heatmaps

The cross-predictability measures illustrated so far capture either the predictive power of one asset's signal for *all* other assets' returns (CPS), or the amount of one asset's return that is predicted by *all* other assets' signals (CPR), or both (CPT), keeping fixed all the rest including assets' own predictability. But how much does a part of the cross-section predict of another specific part, for example what is the value of cross-predicting small-stocks using only large stocks? To answer this question, I propose a more granular cross-prediction measure that sets to zero just one entry of the prediction matrix at a time:

$$\text{value of Cross-prediction Element } (i,j) := CPE^{\bar{R}}_{i,j} = 1 - \frac{\bar{R}_{-ij}}{\bar{R}_{full}} \tag{20}$$

for any $i, j = 1, ..., N$. The measure is defined analogously for SR, which I denote as $CPE^{SR}_{i,j}$. CPE captures the loss in portfolio performance due to the exclusion of specific cross-prediction effects from asset j's signal to asset i's return, and can be applied also to elements of the main diagonal of Π. With $N = 10$ assets, I get $N^2 = 100$ CPEs, which can be arranged into a $N \times N$ matrix similar to a correlation matrix. Figure 8 shows a heatmap for $CPE^{\bar{R}}$ for the first PP, PEP and PAP (from top to bottom) on the left and CPE^{SR} on the right. I name these figures *Cross-Predictability Heatmaps (CP-Heatmaps).* The values are averages over PPA carried out with one of 153 signals at a time, as above. Starting from zeros, which are blank spaces, darker colors represent higher values, according to the colored bar on the right of each panel. Green denotes positive values and red negative ones.

Let us start from the PP case. One region of interest is the main diagonal, whose elements refer to own-asset prediction effects, keeping all other cross-prediction effects captured by PPA fixed. The signal of small stocks positively contributes to predict themselves ($CPE_{1,1}$ is about 0.11). Interestingly, own-asset cross-prediction values are quite low for the rest of the cross-section and become slightly negative for larger stocks. Hence, the CP-heatmap reveals why $\tilde{F}$ underperforms compared to LS in Figure 5. As in the previous sections, cross-predictability concentrates on small firms, and much more so for cross-predicted returns (first row) than for cross-predictive signals (first column). More in detail regarding the first row, the second half of the size-sorted portfolios predict small stocks better than the first half, with CPE increasing up to a peak for the tenth portfolio. The last row shows that the cross-predictive power of large companies is mainly due to predicting small stocks, analogously to the lead-lag relation documented in the literature. Finally, the CP-heatmap is overall greener in the upper portion and more red in the lower portion, which suggests that larger stocks are more difficult to predict. The middle of the heatmap is generally populated with lower values, i.e., cross-predictability concentrates on the extremes.

The second panels of Figure 8 refer to symmetric strategies (PEP). In general, the CP-heatmap looks very similar to the PP case, with the main difference that there are fewer negative values here. Now larger stocks contribute positively to portfolio performance when predicting their own returns (bottom-right element). Differences compared to the top panel originate in the nature of symmetric strategies: although cross-prediction terms appear in Π^s as showed in Section 4, PEPs contain factor risk and are based on own-asset predictions, which have here the highest importance ($CPE_{1,1} = 0.31$ and $CPE_{10,10} = 0.13$).

Antisymmetric strategies are based purely on cross-prediction effects and are factor-neutral. From the bottom panels, this is immediately clear as the CPE along the main diagonal is always perfectly zero, as is the main diagonal in Π^a. Also for PAP, the cross-prediction effects contributing the most to portfolio performance involve small stocks. Preventing large firms from predicting them reduces the average return of the first PAP by 15%. Small companies, instead, have a negative impact when used to predict other assets (first row), and it becomes increasingly more difficult to predict larger stocks: from top to bottom, the CP-heatmap passes from green to red. The CP-heatmap passes instead from red to green from left to right, meaning that signals' cross-predictive power increases with size. CPE^{SR}'s (right-side of Figure 8) show basically the same results.

CP-heatmpas help to further understand the sources of cross-predictability within the PPA framework. Empirical results suggest again that cross-predictability concentrates on small stocks returns predicted using other assets' signals, predominantly large stocks. The amount of returns cross-predicted by other assets tend to decrease with size while assets' cross-predictive power tends to increase with market capitalization.

7 Discussion: Sources of Cross-Predictability

After having explored *how much* cross-predictability there is in the cross-section, I now address the question regarding *what* drives cross-predictability. Previous research has suggested several explanations based on slow information diffusion (Menzly and Ozbas, 2010; Müller, 2019) and frictions to information processing (Cohen and Lou, 2012). Burt and Hrdlicka (2021) argues that own-firm momentum captures the bulk of cross-predictability at longer horizons. As discussed before, these papers mainly rely on long-short portfolios disregarding the rest of the cross-section. With the approach used in this paper, I can investigate what influences cross-predictability with more precise tests. The results of this section are based on the average over all 153 signals considered to ensure that they are not driven by signal-specific behaviors.

To start with, in Figure 9 I plot time-series of $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$ for the ten size-sorted portfolios. I build these by applying Eq.(14) to (16) to rolling average returns instead of standard averages, using 120-month windows (like in the main PPA implementation in KMP) that require at least 95% available observations in each window. Calculating the measures on rolling averages of returns reduces the sensitivity to extreme observations that otherwise would arise by taking ratios of returns in each month directly whenever the denominator (PP return) is close to zero.

The cross-predictability measure for small stocks stands out, being well above any other portfolio for most of the sample for $CPT^{\bar{R}}$ and $CPR^{\bar{R}}$. In the second panel, relative to $CPS^{\bar{R}}$, the largest companies have the highest importance whereas the smallest stocks have the lowest, contributing negatively most of the time. An important change concerning small firms occurs towards 2014, when $CPT^{\bar{R}}$ and $CPR^{\bar{R}}$ drop and $CPS^{\bar{R}}$ increases. The opposite happens especially for the second portfolio. It is difficult to argue whether this trend will remain stable in the future considering the high volatility of the time-series.

What determines cross-predictability? The main finding of the previous sections is that the largest economic value of cross-predictability lies in cross-predicting small stocks through the signals other assets, especially large companies. The contribution of the latter alone, however, is relatively low. More generally, the power of cross-predictive signals increases with market capitalization whereas cross-predicted returns shrink. These facts, paired with the previous literature, hint to an explanation based on information frictions and limits to arbitrage. Typically, small stocks are more difficult to trade than large stocks due to lower liquidity and higher transaction costs, and as such potential arbitrage opportunities involving these assets might not be fully exploited or exploited only with some delay. In general, I expect strongly mispriced assets to be more predictable than less mispriced ones, and not necessarily only through their own signal. For instance, signals originating from the rest of the cross-section might help predicting at least a part of small stocks' returns, but not large stocks' returns. This could explain why $CPR^{\bar{R}}$, which captures cross-predictability *beyond* own-asset predictions, is large for portfolio one but small or even negative for others.

To investigate this conjecture, I use cross-sectional Fama-MacBeth regressions (FMB, Fama and MacBeth (1973)) for the ten size-sorted portfolios:

$$CP_t = \alpha_t + \beta_1 MIS_t + \beta_2 X_t + \varepsilon_t \tag{21}$$

where $CP_t = \{CPT^{\bar{R}}, CPS^{\bar{R}}, CPR^{\bar{R}}\}$, contains one of the three cross-predictability measures for the ten test assets, MIS_t is the $MGMT$ mispricing measure from Stambaugh and Yuan (2017), and X_t is a vector of controls.[16] Since Figure 9 shows that small stocks have a dominant role and their cross-prediction measures are disproportionately more volatile than the rest, I normalize the measure of each portfolio by its standard deviation. Accordingly, I do the same for the variables on the right-hand side.[17]

[16] Stambaugh and Yuan (2017) also provide $PERF$, another mispricing measure based on a set of "performance" factors. Due to its very high correlation with the controls used (e.g. 0.54 with book-to-market), I exclude it from the analysis as the resulting estimates would be impaired by high collinearity problems.

[17] In this section I report results only for PP for reasons of space. Moreover, since PPs represent the complete outcome of PPA, cross-predictability measures are easier to interpret than those based on PEPs or PAPs.

I report the results in Table 2. The panels refer to $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$ from top to bottom, respectively. In the first column, the model contains portfolio-level book-to-market, momentum (Carhart, 1997) and operating profitability (Fama and French, 2015) as controls. The coefficient for MIS is positive (0.25) and highly significant, which means there is a positive correlation between mispricing and cross-predictability. The book-to-market ratio and operating profitability have a strongly negative effect with coefficients of -0.35 and -0.87, respectively, whereas momentum has positive but barely significant impact. The negative effect of book-to-market resembles the finding in Kelly and Pruitt (2013), where a cross-sectional book-to-market measure constructed with Partial Least Squares (PLS) predicts the returns of growth stocks better than value stocks.[18]

In column 2 I add idiosyncratic return volatility with respect to Fama and French (1993) model, $IVOL$ (Ang et al., 2006). Cohen and Lou (2012) suggest that idiosyncratic volatility proxies for limits to arbitrage. *Coeteribus paribus*, stocks that are more difficult to arbitrage are likely to be more mispriced, such that signals coming from other assets can be useful in predicting a part of their future price movements. The coefficient for $IVOL$ is 0.64 and it is significant at any conventional significance level, as I would expect.

The model specification in the last column adds size. Other than being an important control, this covariate is interesting on its own as it can also proxy for limits to arbitrage (Cohen and Lou, 2012). Size has a large negative effect, which means the larger the stocks, the lower their total cross-prediction value, a finding that dovetails with the trend observed above that $CPT^{\bar{R}}$ is large for small stocks but decreases for bigger firms, which are typically

The overall findings do not change substantially compared to PPs. They are available upon request.

[18]See Table III in their paper.

those with the smallest limits to arbitrage.[19] The coefficients of the other controls do not change considerably. The average cross-sectional R^2 of the regressions is roughly 50%.

The second panel reports results for $CPS^{\bar{R}}$. While, in the first specification, MIS has a positive coefficient, this becomes strongly negative in the last two specifications, with a t-statistic above 3 in absolute value. Also $IVOL$ has a negative effect. This opposite behavior compared to $CPT^{\bar{R}}$ makes sense intuitively: if an asset is strongly mispriced or if it is difficult to close arbitrage opportunities connected to it, its signal becomes less reliable and loses cross-predictive power, and since $CPR^{\bar{R}}$ dominates $CPS^{\bar{R}}$, the overall effect has opposite sign compared to $CPT^{\bar{R}}$. In similar fashion, the sign of size also changes, but the coefficient is not significant. Regarding the controls, book-to-market and operating profitability increase the signal cross-predictive power like momentum, which is now strongly significant. Comparing the second to the third panel, which refers to $CPR^{\bar{R}}$, the two components of the total cross-predictability value are almost opposed, with regression coefficients flipping sign. Nevertheless, the explanatory power (average cross-sectional R^2) is very similar in all the three cases, which suggests that such variables have a strong and stable influence on assets' cross-prediction measures.

In Table 3, I repeat the same cross-sectional regressions for cross-predictability measures based on SRs. The results remain vastly unchanged: MIS, $IVOL$ and Size have strongly significant coefficients with the same sign as before (with the only exception that MIS is insignificant in the last model specification for CPS^{SR}), and the sign of the independent variables tends to flip between CPR^{SR} and CPS^{SR}. The cross-sectional R^2 is some percentage

[19] Kelly and Pruitt (2013) predict better large stocks than small stocks, instead. The reader should bear in mind that their PLS-based book-to-market represents the entire cross-section and not only one portion like here. Moreover, it cannot distinguish between cross-predictive signals and cross-predictive returns as my measures. Nonetheless, it represents a meaningful comparison for the results of this paper.

points lower compared to Table 2.

In this section I documented results that help shedding light concerning the determinants of cross-predictability, which seem to be mostly represented by limits to arbitrage and mispricing. These findings can be a useful starting point for future research aimed at building a formal structural model that creates cross-prediction patterns, an approach that has not been adopted yet in the literature.

8 Robustness Tests: Double-sorted Portfolios

A natural extension to single-sorted portfolios are double-sorted portfolios. Two-sorts have been extensively employed as a non-parametric tool to control for the effect of one covariate. Portfolios sorted along the second dimension (e.g. book-to-market) within a certain range of variation of the first one (e.g. bottom 20% of the size distribution) reflect variation in the second covariate keeping fixed the value of the other one. The 25 size-and-book-to-market-sorted portfolios are among the most used test assets in empirical asset pricing, and I employ them here as a robustness test to verify that the findings found above continue to hold with new assets. After all, Section 7 reveals that, other than size, book-to-market is associated with cross-predictability. Portfolios' returns and signals are built with the same approach as for the ten size-sorted portfolios, with the difference that stocks are independently sorted into 5×5 groups by size and by book-to-market ratio. Independent sorts are widely used, such as in Fama and French (1992) or, more recently, in Huang et al. (2022). I apply PPA on the double-sorted portfolios one signal at a time over the 153 signals available and present average results consistently with Section 6.2.

To start with, Figure D.1 in Appendix D shows SRs of the first PP, PEP and PAP together with the simple factor $\tilde{F}$ and the LS factor. To mimick Fama-French size factors, the latter is built going long an equally-weighted average of the 5 bottom size-portfolios and by shorting an equally-weighted average of the 5 top size-portfolios. All three PPA-based strategies abundantly outperform both $\tilde{F}$ and LS, with SRs of about 0.8 for both PP and PEP and above 1 for PAP. Hence, PPA proves very powerful also for double-sorted portfolios.

Moving to cross-predictability measures, Figure D.2 in Appendix D shows $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ for the first PP, PEP and PAP from top to bottom, respectively.[20] Starting from PPs, $CPT^{\bar{R}}$ is remarkably high for the first size quintile (the first five portfolios) and tends to decrease towards the right, i.e., with size, as with single-sorted portfolios. Within each size quintile, $CPT^{\bar{R}}$ generally decreases from left to right with book-to-market, which reflects the negative coefficient of this characteristic in the previous FMB regressions. An exception to this is the first size quintile, where $CPT^{\bar{R}}$ peaks for small-growth stocks ($CPT^{\bar{R}}_{11} = 0.30$), decreases but it is again large for small-value stocks ($CPT^{\bar{R}}_{15} = 0.20$). Overall, the fact that there is low value in the middle of the cross-section occurs relatively clearly within each size quintile. As previously observed, the strongest component of $CPT^{\bar{R}}$ is $CPR^{\bar{R}}$ for smaller stocks, while the opposite holds for larger firms, i.e., the cross-predictive power of large stocks dominates ($CPS^{\bar{R}}$) the amount of their returns that can be predicted by other assets' signals ($CPR^{\bar{R}}$), which becomes even slightly negative on the right of the picture.

Cross-predictability is the highest for small companies also for symmetric strategies (middle panel) and, even if less clearly compared to above, it decreases to zero or even negative

[20] For reasons of space, the first digit identifies the portfolio number along the size dimension and the second one along the book-to-market dimension. For example, portfolio "14" indicates stocks belonging to the first size quintile and the fourth book-to-market quintile.

values in the middle of the cross-section and increases again for larger stocks. The third panel (PAP) is extremely similar to the PP case, both in terms of patterns and magnitudes, pointing again at the fact that the antisymmetric component of PPA-based trading strategies dominates the symmetric component. These findings hold also when using cross-predictability measures based on SRs, illustrated in Figure D.3.

To conclude, I repeat FMB regressions according to Eq.(21) using double-sorted portfolios after building time-series of cross-predictability measures. Results are shown in Table D.1 in Appendix D for $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$. Starting from $CPT^{\bar{R}}$, MIS has now a negative and strongly significant coefficient opposed to the positive coefficient for size-sorted portfolios. Although counter-intuitive at first, one must acknowledge that the magnitude of the coefficient is very low 0.03, roughly ten times lower than for the size-sorted case). Moreover, the coefficient is not significant anymore in the $CPR^{\bar{R}}$ case (third panel). Hence, the relation between cross-predictability and mispricing is feeble for double-sorted portfolios. What remains quite strong, instead, is the impact of idiosyncratic volatility (0.12, t-statistic of 2.76 in column 3) and size (-1.17, t-statistic of -2.93): the stronger the limits to arbitrage, the higher the cross-predictability. Among the controls, momentum and operating profitability are generally insignificant whereas book-to-market ratio has a strongly negative coefficient (-0.34, t-statistic of -8.7), conforming with Figure D.2. Analogously to size-sorted portfolios, on the one hand these results for $CPR^{\bar{R}}$ are almost the same; on the other hand almost all coefficient change sign for $CPS^{\bar{R}}$ (second panel) apart from book-to-market. Said differently, signals' cross-predictive power decreases with mispricing and limits to arbitrage and increases with size, operating profitability and momentum.

The findings are virtually unchanged using SR-based cross-predictive measures (Table

D.2). Notably, here MIS does not significantly impact neither CPT^{SR} nor CPR^{SR}, and $IVOL$ and size remain crucial determinants of cross-predictability.

To sum up, the conclusions relative to cross-predictability reached before continue to hold also with double-sorted portfolios. Most of the value of cross-predictability concentrates on small stocks (specifically small-growth and small-value stocks) and decreases with size and value. Companies with higher market capitalization and book-to-market ratios contribute to PPA-based strategies more by cross-predicting other stocks than vice-versa, in contrast to small stocks. Limits to arbitrage increase the amount of cross-predicted returns but reduce the power of cross-predictive signals.

9 Conclusion

In this paper, I propose a measure based on portfolio performance to quantify the value of cross-predictability. Within the PPA framework introduced by KMP, imposing restrictions on the prediction matrix allows to isolate the impact of cross-prediction effects on optimal linear strategies and to distinguish between the predictive power of one asset's signal for other assets' returns (cross-predictive signals) from the amount of one asset's returns predicted by other assets' signals (cross-predicted returns). my approach gives a comprehensive view on cross-prediction effects on the entire cross-section, differently from the existing practice in the literature that focuses mainly on the extremes (e.g. lead-lag effects).

Using ten size-sorted portfolios as test assets, I observe that most of the value of cross-predictability arises from predicting small stocks' returns with other assets, especially larger companies. In other words, failing to account for cross-prediction effects involving these assets can severely impair the investment outcome. In fact, imposing restrictions on cross-prediction effects regarding small firms triggers a considerable rebalancing in the optimal portfolio weights. Overall, the higher the market capitalization, the more difficult it is to cross-predict stocks' returns. The cross-predictive power of the same asset instead tends to rise with size, although its value is rarely more important than the value of cross-predicted returns. The middle of the cross-section does not significantly contribute to PPA outcomes through cross-prediction effects.

I investigate the link between cross-preditability measures and classical firm characteristics, finding a strong positive relation with mispricing and with limits to arbitrage, and a negative relation with value and profitability. These relations flip sign for cross-predictive signals. Put differently, the more stocks are subject to limits to arbitrage, the more they are mispriced and thus the more room there is for overall predictability, such that other assets' signals can significantly contribute to investment outcomes if employed to predict at least a portion of those stocks' returns. The opposite holds instead when focusing on the cross-predictive power of one asset for other assets' returns: the less the mispricing and limits to arbitrage, the cleaner the signal becomes to predict other assets.

This paper offers a straightforward yet effective methodology to quantify and isolate the sources of cross-predictability in the cross-section. Cross-predictability (including lead-lag effects) has been under the scrutiny of many scholars so far, each proposing different interpretations that frequently exhibit gaps or contradictions. my findings stand as a starting point for future research to facilitate the development of a comprehensive framework that systematically explains all the stylized facts documented in the literature.

References

Ali, U., and Hirshleifer, D. 2020. Shared analyst coverage: Unifying momentum spillover effects. *Journal of Financial Economics* 136:649–675.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. 2006. The cross-section of volatility and expected returns. *Journal of Finance* 61:259–299.

Badrinath, S. G., Kale, J. R., and Noe, T. H. 1995. Of shepherds, sheep, and the cross-autocorrelations in equity returns. *Review of Financial Studies* 8:401–430.

Bagnara, M. 2022. Asset Pricing and Machine Learning: A critical review. *Journal of Economic Surveys,* pp. 1–30.

Brennan, M. J., Jegadeesh, N., and Swaminathan, B. 1993. Investment analysis and the adjustment of stock prices to common information. *Review of Financial Studies* 6:799–824.

Burt, A., and Hrdlicka, C. 2021. Where does the predictability from sorting on returns of economically linked firms come from? *Journal of Financial and Quantitative Analysis* 56:2634–2658.

Campbell, J. Y., and Thompson, S. B. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509–1531.

Carhart, M. M. 1997. On persistence in mutual fund performance. *Journal of Finance* 52:57–82.

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

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

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

Croce, M. M., Marchuk, T., and Schlag, C. 2023. The leading premium. *Review of Financial Studies* 36:2997–3033.

Detzel, A., and Strauss, J. 2018. Combination return forecasts and portfolio allocation with the cross-section of book-to-market ratios. *Review of Finance* 22:1949–1973.

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

Fama, E. F., and French, K. R. 1992. The cross-section of expected stock returns. *Journal of Finance* 47:427–465.

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

Fama, E. F., and MacBeth, J. D. 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 81:607–636.

Freyberger, J., Neuhierl, A., and Weber, M. 2020. Dissecting characteristics nonparametrically. *Review of Financial Studies* 33:2326–2377.

Gibbons, M. R., Ross, S. A., and Shanken, J. 1989. A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society* pp. 1121–1152.

Giglio, S., Kelly, B., and Xiu, D. 2022. Factor models, machine learning, and asset pricing. *Annual Review of Financial Economics* 14:337–368.

Gofman, M., Segal, G., and Wu, Y. 2020. Production networks and stock returns: The role of vertical creative destruction. *Review of Financial Studies* 33:5856–5905.

Goodarzi, M., Schlag, C., and von den Hoff, S. 2022. A New Model Every Month? — Dynamic Model Selection for Stock Return Prediction. Working paper.

Greene, W. H. 2003. *Econometric analysis*. Pearson Education India.

Gu, S., Kelly, B., and Xiu, D. 2020. Empirical asset pricing via machine learning. *Review of Financial Studies* 33:2223–2273.

Hou, K. 2007. Industry information diffusion and the lead-lag effect in stock returns. *Review of Financial Studies* 20:1113–1138.

Hou, K., Xue, C., and Zhang, L. 2020. Replicating anomalies. *Review of Financial Studies* 33:2019–2133.

Huang, S., Lee, C. M., Song, Y., and Xiang, H. 2022. A frog in every pan: Information discreteness and the lead-lag returns puzzle. *Journal of Financial Economics* 145:83–102.

Jensen, T. I., Kelly, B., and Pedersen, L. H. 2022. Is there a replication crisis in finance? *The Journal of Finance,* Forthcoming.

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

Kelly, B., and Pruitt, S. 2013. Market expectations in the cross-section of present values. *Journal of Finance* 68:1721–1756.

Kelly, B. T., Malamud, S., and Zhou, K. 2022. The virtue of complexity in return prediction. Tech. rep., National Bureau of Economic Research.

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

Kozak, S., Nagel, S., and Santosh, S. 2018. Interpreting factor models. *Journal of Finance* 73:1183–1223.

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

Lee, C. M., Sun, S. T., Wang, R., and Zhang, R. 2019. Technological links and predictable returns. *Journal of Financial Economics* 132:76–96.

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

Lo, A. W. 2002. The statistics of Sharpe ratios. *Financial analysts journal* 58:36–52.

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

Menzly, L., and Ozbas, O. 2010. Market segmentation and cross-predictability of returns. *Journal of Finance* 65:1555–1580.

Müller, S. 2019. Economic links and cross-predictability of stock returns: Evidence from characteristic-based "styles". *Review of Finance* 23:363–395.

Newey, W. K., and West, K. D. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55:703–708.

Parsons, C. A., Sabbatucci, R., and Titman, S. 2020. Geographic lead-lag effects. *Review of Financial Studies* 33:4721–4770.

Rapach, D. E., Strauss, J. K., and Zhou, G. 2013. International stock return predictability: what is the role of the United States? *Journal of Finance* 68:1633–1662.

Schlag, C., and Zeng, K. 2019. Horizontal industry relationships and return predictability. *Journal of Empirical Finance* 53:310–330.

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

Figures

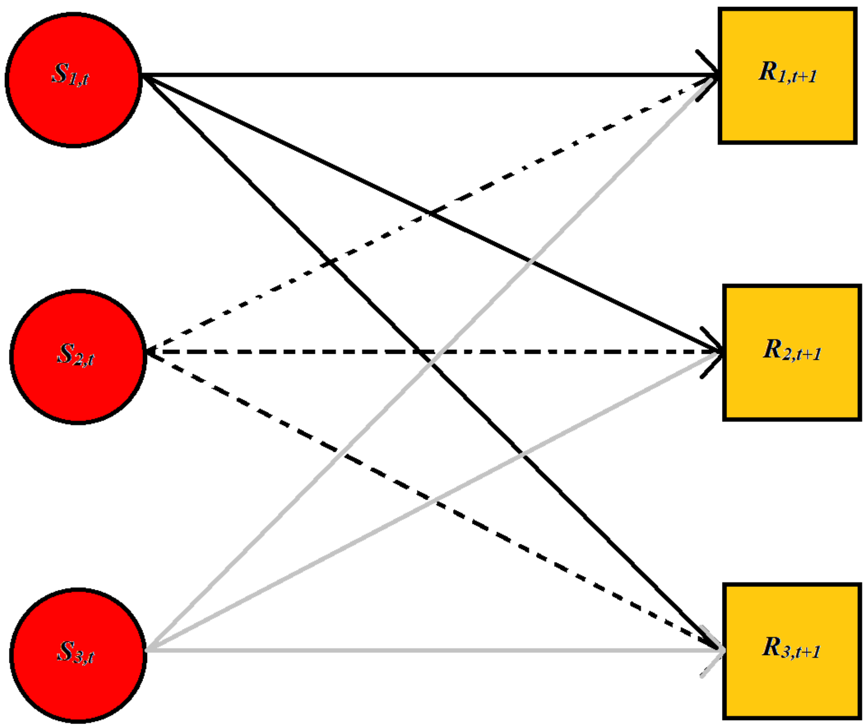

FIGURE 1: **Example of cross-prediction effects**

This figure shows how cross-prediction effects can arise in a simple case with $N = 3$ stocks with a signal $S_t = (S_{1,t}, S_{2,t}, S_{3,t})$ (red circles) used to predict the returns $R_{t+1} = (R_{1,t+1}, R_{2,t+1}, R_{3,t+1})$ (yellow squares). Horizontal arrows represent own-asset predictions. Arrows originating from each $S_{j,t},\ j = 1, 2, 3$ to other assets' return $R_{j,t+1},\ j \neq i$ capture cross-predictive signals for asset j. For asset $j,\ 1, 2, 3$, arrows pointing towards $R_{j,t+1}$ and originating from other assets' signals $S_{j,t},\ j \neq i$ illustrate cross-predicted returns. The cross-prediction total for asset j is given by putting together its cross-predictive signal and its cross-predicted return. Arrows originating from each signal have the same style, such as a dashed line.

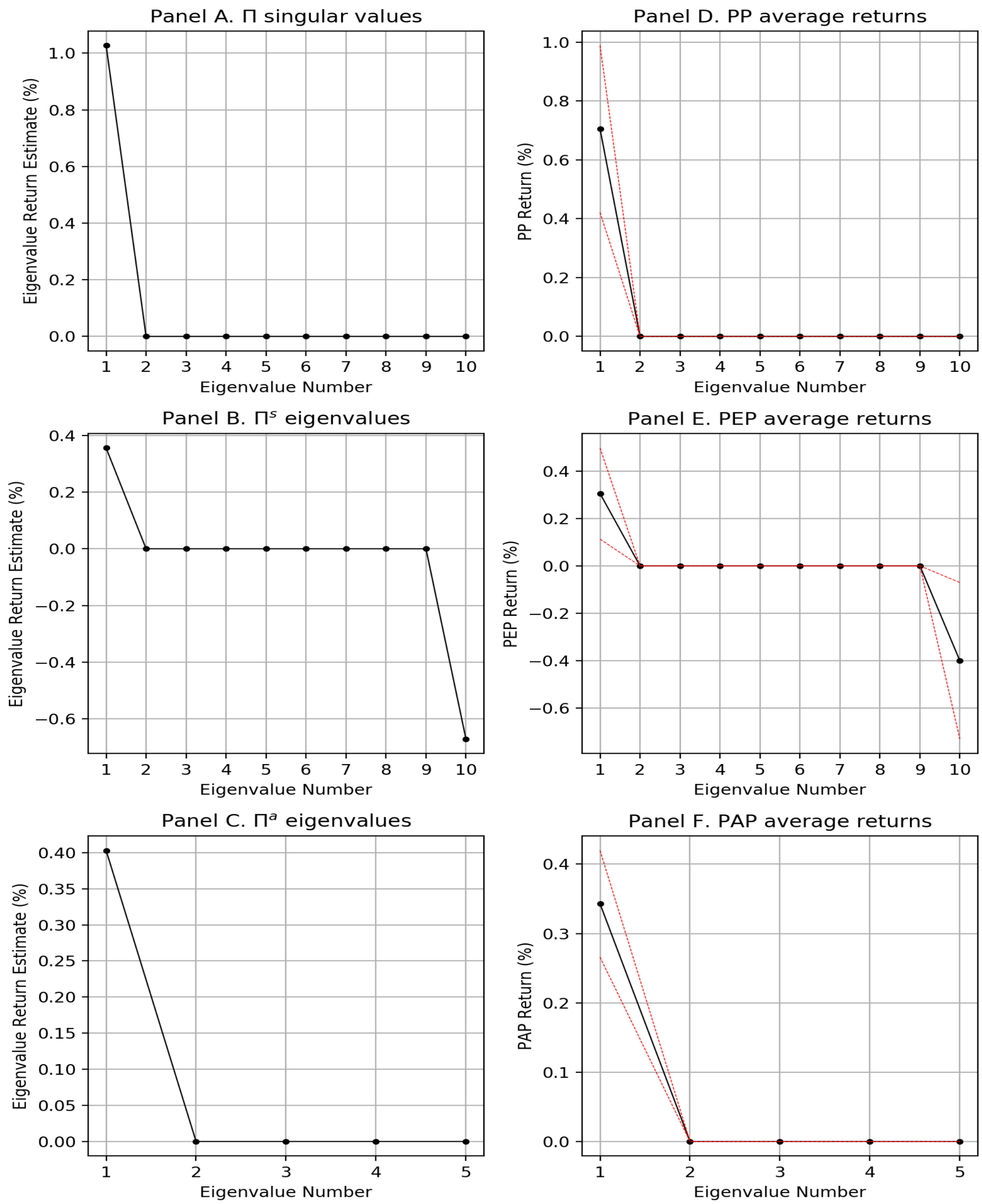

FIGURE 2: **Prediction matrix eigenvalues**

In this figure, panels A, B, and C show estimated singular values and eigenvalues for Π, Π^s and Π^a averaged over training samples for ten size-sorted portfolios (portfolio number on x-axis). Panels D, E, and F show average out-of-sample returns and ± 2 standard error confidence bands for corresponding PPs, PEPs and PAPs, respectively. The sample period is 1968 - 2019.

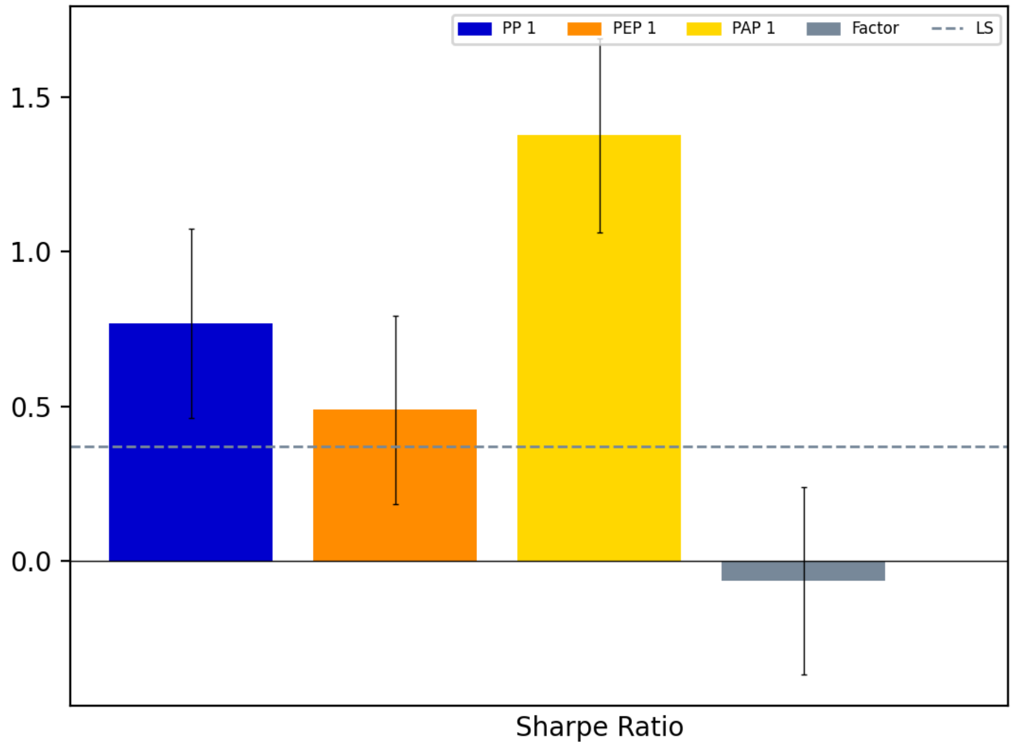

FIGURE 3: **PPA Performance**

This figure shows out-of-sample annualized SRs for the first PP, PEP and PAP, along with ±2 standard error bands around each estimate, for ten size-sorted portfolios. PPA is carried out using size as signal. Each forecast is made on an out-of-sample basis using a rolling training window of 120 months. The sample period is 1968 - 2019.

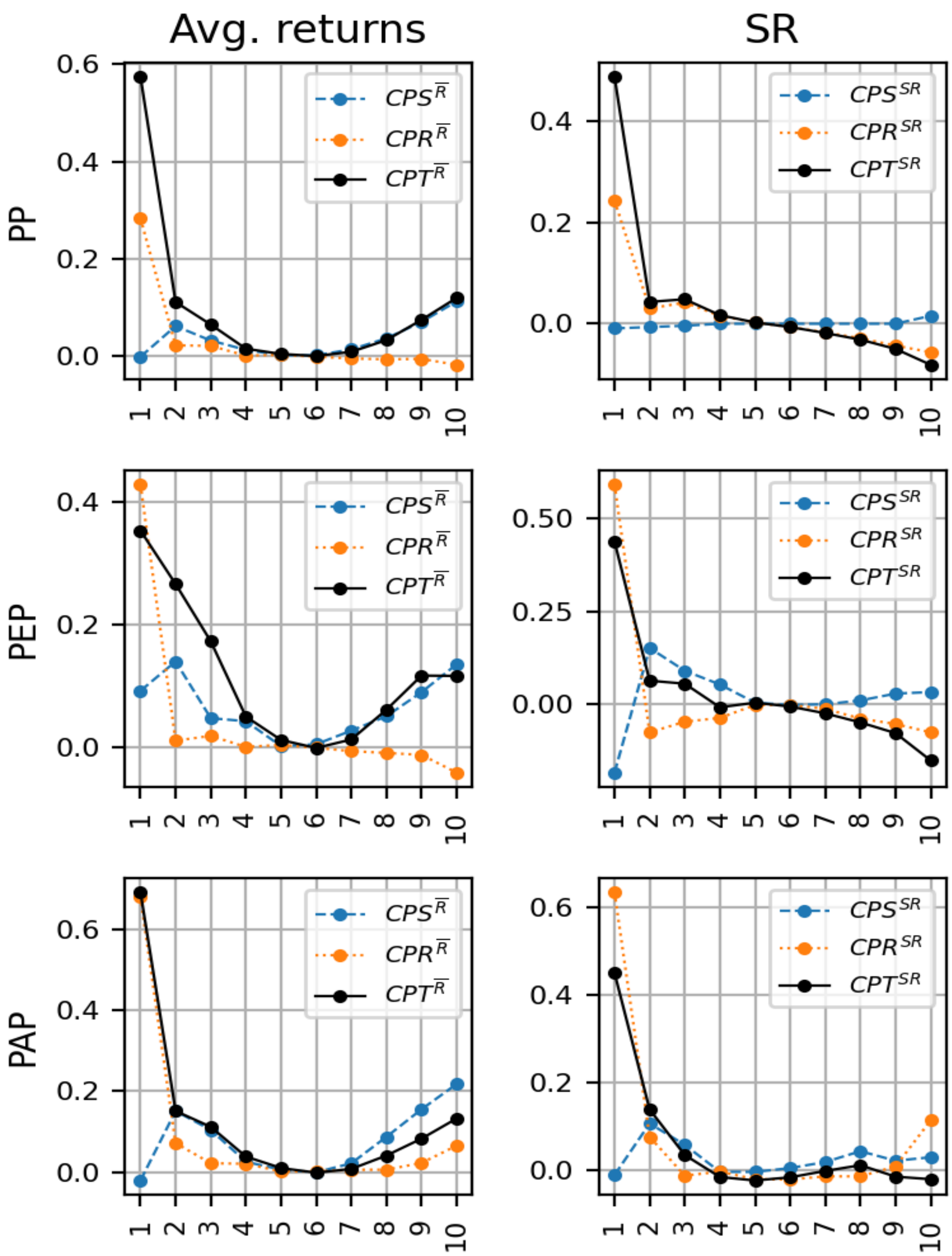

FIGURE 4: **Size-sorted Portfolios:** $CPS^{\bar{R}}$**,** $CPR^{\bar{R}}$ **and** $CPT^{\bar{R}}$ **(**CPS^{SR}**,** CPR^{SR} **and** CPT^{sr}**)**
This figure shows $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ on the left and CPS^{SR}, CPR^{SR} and CPT^{SR} on the right for ten size-sorted portfolios (portfolio number on x-axis) where PPA is carried out using size as signal. $\bar{R}$ refers to average realized returns for a trading strategy based on the first PP, PEP and PAP, from top to bottom. Analogously for SR. The sample period is 1968 - 2019.

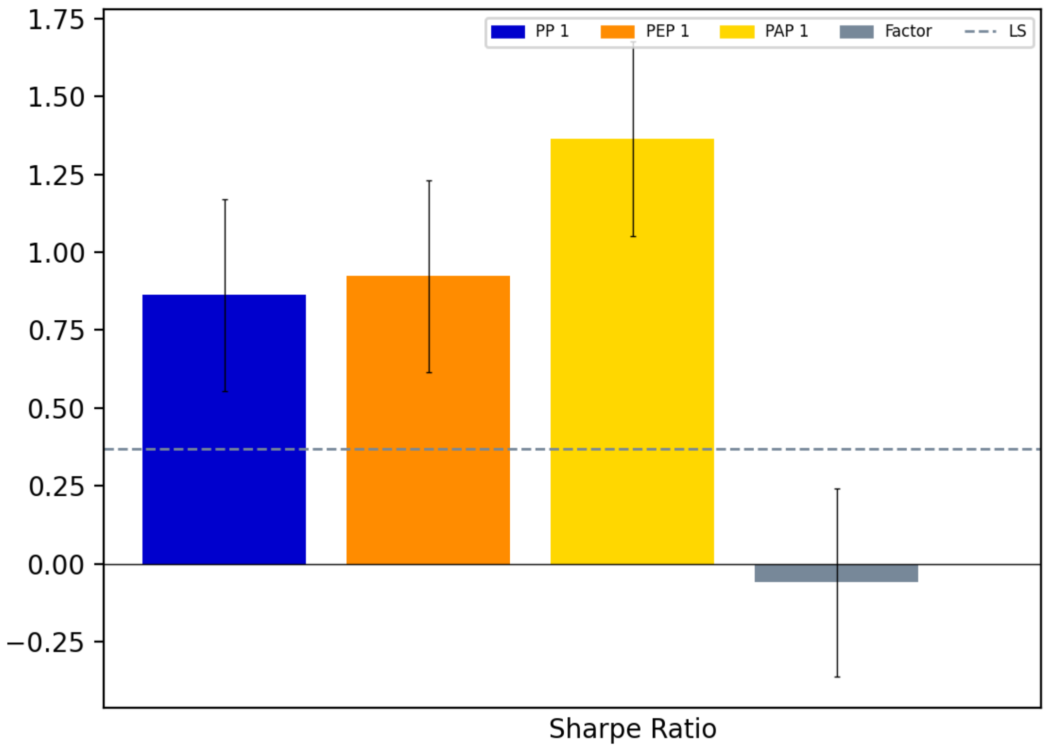

FIGURE 5: **Performance ratios across all signals**

This figure shows average out-of-sample annualized SRs for the first PP, PEP and PAP over 153 signals for ten size-sorted portfolios, along with ±2 standard error bands around each estimate. PPA is carried out using one signal at a time. Each forecast is made on an out-of-sample basis using a rolling training window of 120 months. The sample period is 1968 - 2019.

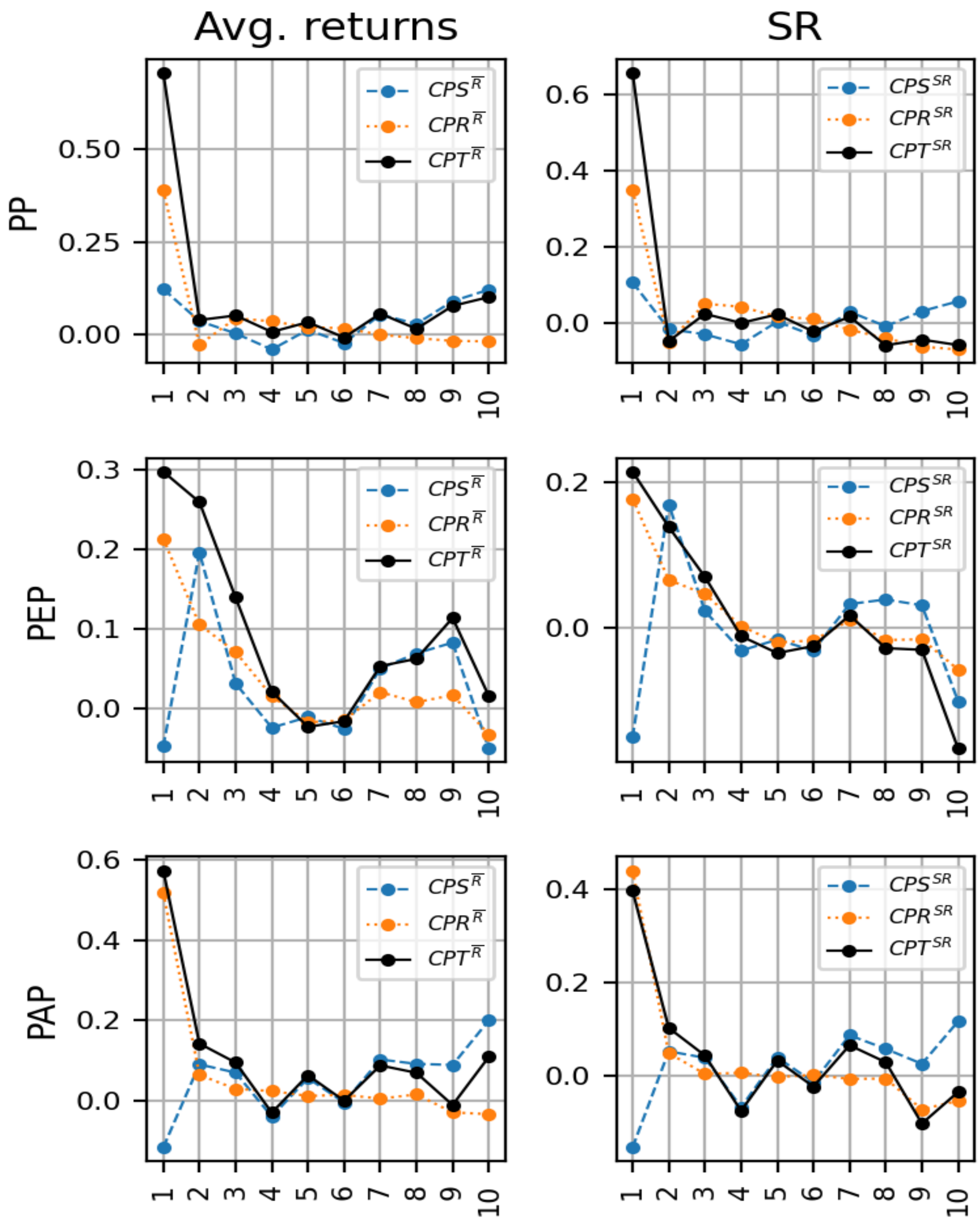

FIGURE 6: **Size-sorted Portfolios:** $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ **and** $CPT^{\bar{R}}$ (CPS^{SR}, CPR^{SR} **and** CPT^{SR}) **across all signals**

This figure shows the average $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ on the left and CPS^{SR}, CPR^{SR} and CPT^{SR} for ten size-sorted portfolios (portfolio number on x-axis) over 153 signals, where PPA is carried out using one signal at a time. $\bar{R}$ refers to average realized returns of a trading strategy based on the first PP, PEP, and PAP, from top to bottom. Analogously for SR. The sample period is 1968 - 2019.

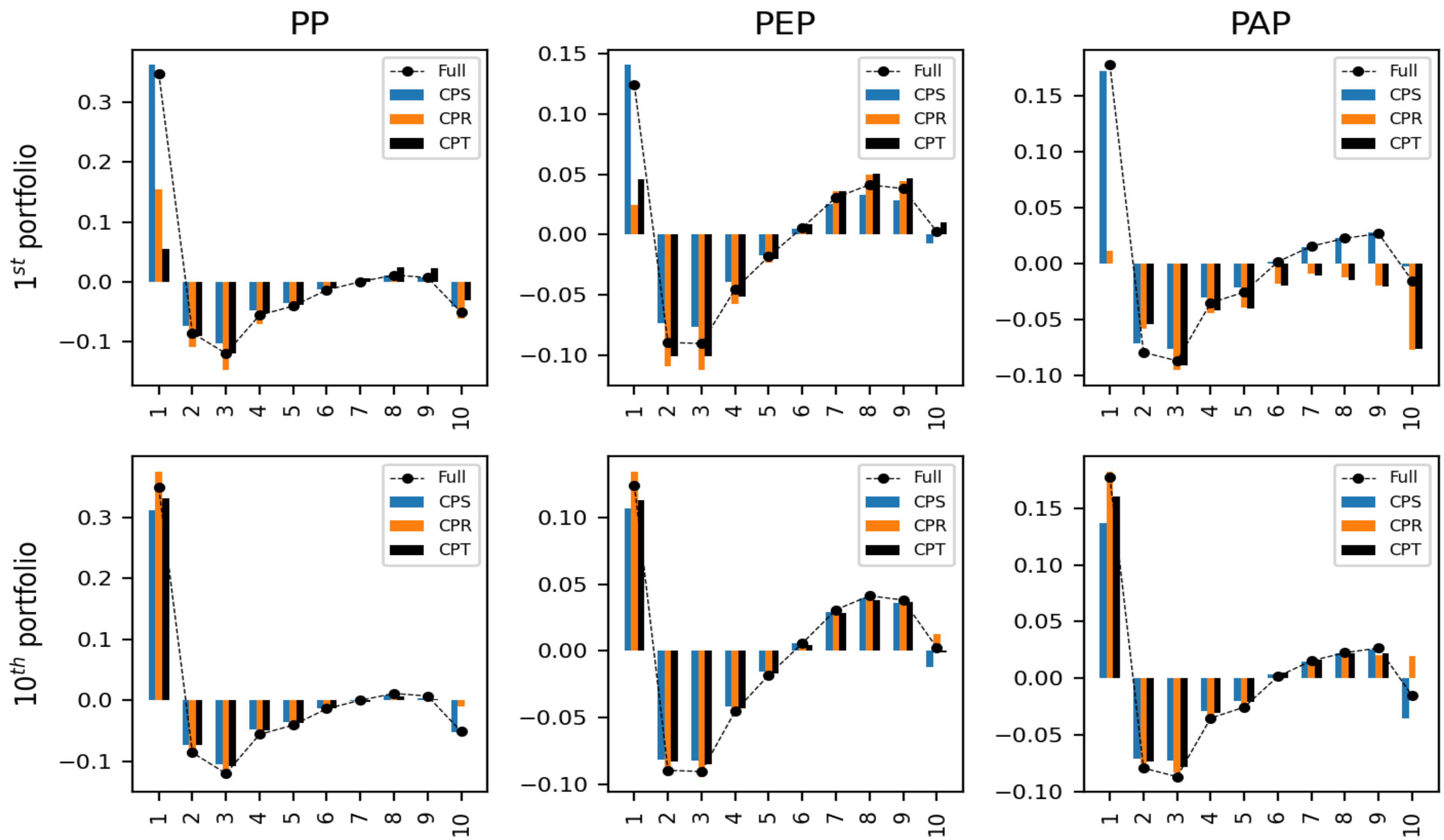

FIGURE 7: **Size-sorted Portfolios: portfolio weights for the first PP**

This figure shows time-series averages of portfolio weights for ten size-sorted portfolios (portfolio number on x-axis) for the first PP, PEP and PAP from left to right, averaged over 153 signals. The dashed line shows weights when the Π matrix is unrestricted. The top (bottom) panels of each column show weights when the cross-predictive signals (CPS), cross-predicted returns (CPR) or cross-prediction total (CPT) corresponding to the first (last) size portfolio are excluded, respectively. The sample period is 1968 - 2019.

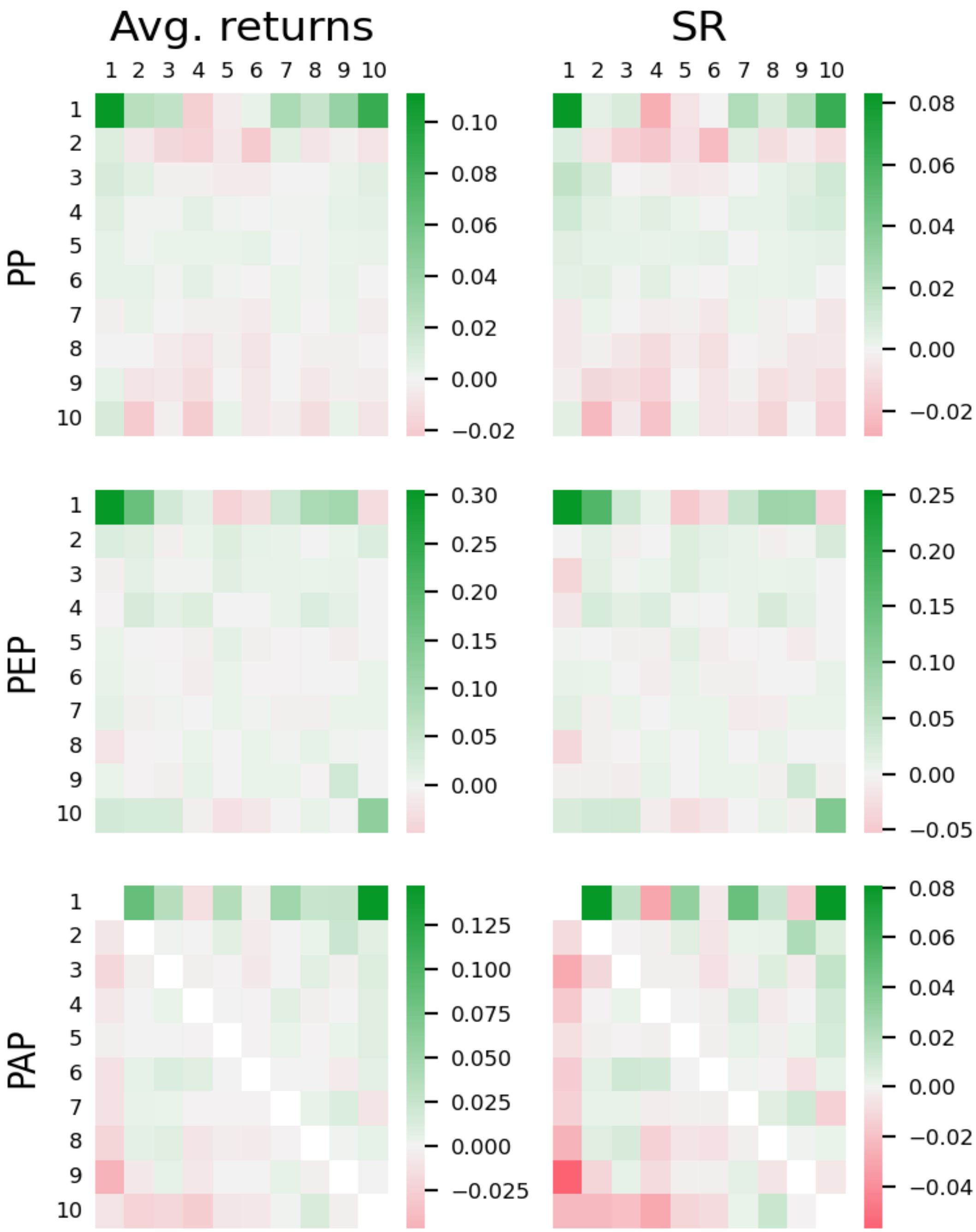

FIGURE 8: **Size-sorted Portfolios:** CEV^R **and** CEV^{SR}

This figure shows CP-heatmaps for average $CPE^{\bar{R}}$ (on the left) and CPE^{SR} (on the right) over 153 signals for ten size-sorted portfolios where PPA is carried out using one signal at a time. From top to bottom, panels represent results for PP, PEP and PAP. Green colors represent positive values and red colors negative ones. The sample period is 1968 - 2019.

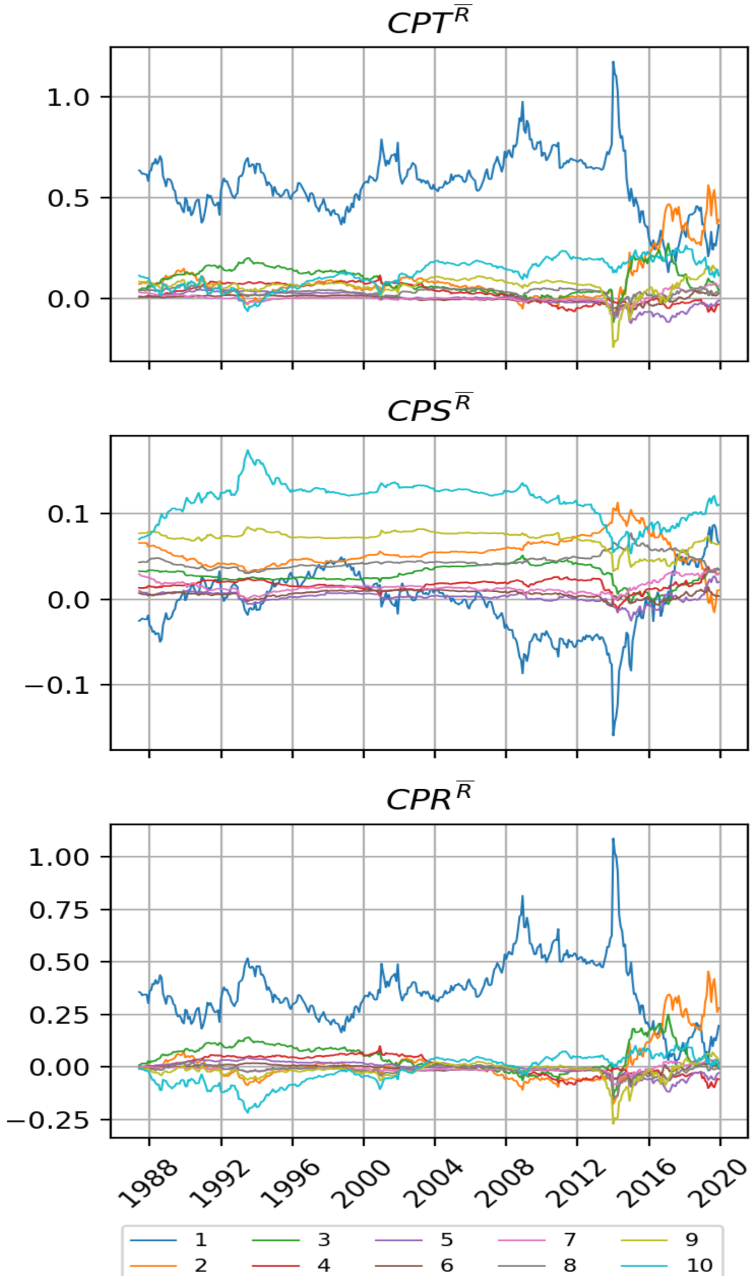

FIGURE 9: **Size-sorted Portfolios: rolling $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ across all signals**

This figure shows rolling estimates of $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$ for ten size-sorted portfolios (portfolio number in legend) averaged over 153 signals, where PPA is carried out using one signal at a time. Rolling average returns of the first PP are computed with a rolling window of 120 months, and the cross-predictability measures are calculated using Eq.(14) to (16). Accounting for the rolling estimation, the available sample period is 1989 - 2019.

Tables

TABLE 1: **Portfolio correlation and explanatory power**

This table reports the correlation matrix for LS, $\tilde{F}$, the first PP, PEP and PAP obtained through PPA analysis on ten size-sorted portfolios. The last row reports the average R^2 from regressing each size-sorted portfolio on a one-factor model corresponding to the factor reported in each column. The sample period is 1968 - 2019.

	LS	$\tilde{F}$	PP	PEP	PAP
LS	1.00	-0.97	0.57	-0.74	0.51
$\tilde{F}$	-0.97	1.00	-0.55	0.79	-0.41
PP	0.57	-0.55	1.00	0.08	0.58
PEP	-0.74	0.79	0.08	1.00	-0.06
PAP	0.51	-0.41	0.58	-0.06	1.00
R^2	0.18	0.14	0.04	0.08	0.02

Table 2: **Cross-predictability: Fama-MacBeth Regressions (Average Returns)**

This table reports the results of Fama-MacBeth (Fama and MacBeth, 1973) cross-sectional regressions according to Eq.(21). The first, second and third panel refer to $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$ based on rolling average returns for ten size-sorted portfolios as dependent variable, respectively. MIS_t is the $MGMT$ mispricing measure from Stambaugh and Yuan (2017) and X_t is a vector of controls which includes portfolio-level book-to-market ratio (BM), momentum (MOM) and operating profitability (OP) in column (1), together with idiosyncratic volatility ($IVOL$) in column (2) and idiosyncratic volatility and Size in column (3). Both dependent and independent variables are normalized by their standard deviation. t-statistics based on Newey-West standard errors (Newey and West, 1987) with 4 lags (chosen following Greene (2003)) are in brackets. Bold numbers indicate significant statistics based on conventional significance levels. The last column of each panel reports the average cross-sectional R^2. Accounting for the rolling estimation, the available sample period is 1989 - 2019.

$CPT^{\bar{R}}$	(1)	(2)	(3)
MIS	**0.25 (17.84)**	**0.32 (18.83)**	**0.35 (14.57)**
$IVOL$		**0.64 (4.32)**	**0.55 (3.54)**
$Size$			**-5.76 (-5.11)**
BM	**-0.35 (-5.3)**	**-0.27 (-3.95)**	**-0.59 (-4.95)**
MOM	0.21 (1.19)	0.1 (0.5)	0.18 (0.9)
OP	**-0.87 (-8.39)**	**-0.89 (-8.63)**	**-0.84 (-6.89)**
R^2	0.51	0.51	0.57
$CPS^{\bar{R}}$	(1)	(2)	(3)
MIS	**0.04 (2.5)**	**-0.11 (-3.72)**	**-0.14 (-3.38)**
$IVOL$		**-1.67 (-5.34)**	**-1.46 (-5.1)**
$Size$			-0.16 (-0.08)
BM	**0.4 (2.92)**	**0.27 (2.02)**	**0.42 (1.69)**
MOM	**2.11 (4.23)**	**2.42 (4.73)**	**2.81 (4.66)**
OP	**0.68 (3.97)**	**0.56 (3.37)**	**0.69 (3.78)**
R^2	0.55	0.56	0.58
$CPR^{\bar{R}}$	(1)	(2)	(3)
MIS	**0.49 (7.42)**	**0.22 (13.23)**	**0.25 (10.81)**
$IVOL$		**0.97 (7.06)**	**0.84 (5.81)**
$Size$			**-4.98 (-5.66)**
BM	**-2.52 (-8.26)**	**-0.55 (-7.51)**	**-0.91 (-7.55)**
MOM	-0.97 (-1.13)	-0.23 (-1.19)	-0.2 (-1.06)
OP	**-2.06 (-4.48)**	**-0.54 (-5.24)**	**-0.5 (-3.97)**
R^2	0.48	0.58	0.62

Table 3: **Cross-predictability: Fama-MacBeth Regressions (SRs)**

This table reports the results of Fama-MacBeth (Fama and MacBeth, 1973) cross-sectional regressions according to Eq.(21). The first, second and third panel refer to CPT^{SR}, CPS^{SR} and CPR^{SR} based on rolling average SRs for ten size-sorted portfolios as dependent variable, respectively. MIS_t is the $MGMT$ mispricing measure from Stambaugh and Yuan (2017) and X_t is a vector of controls which includes portfolio-level book-to-market ratio (BM), momentum (MOM) and operating profitability (OP) in column (1), together with idiosyncratic volatility ($IVOL$) in column (2) and idiosyncratic volatility and Size in column (3). Both dependent and independent variables are normalized by their standard deviation. t-statistics based on Newey-West standard errors (Newey and West, 1987) with 4 lags (chosen following Greene (2003)) are in brackets. Bold numbers indicate significant statistics based on conventional significance levels. The last column of each panel reports the average cross-sectional R^2. Accounting for the rolling estimation, the available sample period is 1989 - 2019.

CPT^{SR}	(1)	(2)	(3)
MIS	**0.04 (4.74)**	**0.16 (10.58)**	**0.18 (7.91)**
$IVOL$		**1.1 (8)**	**1.16 (7.75)**
$Size$			**-2.4 (-2.69)**
BM	**-0.82 (-11.18)**	**-0.75 (-9.65)**	**-0.91 (-8.66)**
MOM	**0.51 (2.44)**	0.3 (1.39)	0.25 (1.1)
OP	**-0.59 (-5.47)**	**-0.58 (-5.48)**	**-0.57 (-4.65)**
R^2	0.44	0.46	0.42
CPS^{SR}	(1)	(2)	(3)
MIS	**0.04 (4.13)**	**0.07 (4)**	0.02 (0.79)
$IVOL$		0.25 (1.34)	**0.42 (2.42)**
$Size$			**5.41 (5.12)**
BM	0.03 (0.39)	0.05 (0.51)	**0.34 (2.58)**
MOM	**0.86 (3.39)**	**0.83 (3.04)**	**1.04 (3.36)**
OP	**-0.4 (-3.78)**	**-0.41 (-3.89)**	**-0.43 (-4.42)**
R^2	0.22	0.22	0.22
CPR^{SR}	(1)	(2)	(3)
MIS	**0.04 (4.06)**	**0.15 (9.5)**	**0.17 (7.61)**
$IVOL$		**0.99 (7.27)**	**0.99 (6.82)**
$Size$			**-3.35 (-3.81)**
BM	**-0.81 (-10.44)**	**-0.75 (-9.48)**	**-0.99 (-8.69)**
MOM	**0.43 (1.92)**	0.25 (1.09)	0.22 (0.9)
OP	**-0.68 (-5.61)**	**-0.67 (-5.62)**	**-0.68 (-5.07)**
R^2	0.47	0.47	0.43

Appendix A Principal Portfolios: Further Mathematical Details

A.1 Bounds on Portfolio Size

The constraint to the problem in (7) is $||L|| \leq 1$. Since it holds that $||L'S_t|| \leq ||L'||\ ||S_t|| \leq ||S_t||$ if $||L|| \leq 1$, the constraint represents a bound on the portfolio size $||L'S_t||$ corresponding to portfolio weights $S_t'L$ that admits only linear strategies with a position size not exceeding the position size of the simple factor $\tilde{F}$. Moreover, if S_t is normalized such that $||S_t|| = 1$ for all signals as in my empirical analysis, only strategies with $||L'S_t|| \leq 1$ are considered.

A.2 Expected Return and Factor Exposure

Given that linear trading strategy can be decomposed into

$$R^w_{t+1} = S_t'LR_{t+1} = S_t'L^sR_{t+1} + S_t'L^aR_{t+1} \tag{A.1}$$

its expected returns is

$$E[R^w_{t+1}] = tr((L^s + L^a)(\Pi^s + \Pi^a)) = tr(L^s\Pi^s) + tr(L^a\Pi^a) \tag{A.2}$$

because $tr(BA) = tr(AB) = 0$ for any symmetric matrix $B \in \mathbb{R}^N$ and any antisymmetric matrix $A \in \mathbb{R}^N$.

Consider a "latent" factor F introduced in KMP that satisfies

$$F_{t+1} = \left(\frac{1}{S_t'(\Sigma_R)^{-1}S_t} (\Sigma_R)^{-1} S_t \right)' R_{t+1} \quad \text{(A.3)}$$

with Σ_R being the return covariance matrix.[1] Since

$$Cov_t(R^w_{t+1}, F_{t+1}) = w_t' Cov_t(R_{t+1}, F_{t+1}) = Var_t(F_{t+1}) S_t' L^s S_t \quad \text{(A.4)}$$

the conditional latent factor exposure of a linear trading strategy is

$$\underbrace{\frac{cov_t(R^w_{t+1}, F_{t+1})}{var_t(F_{t+1})}}_{factor\ beta} = S_t' L^s S_t \quad \text{(A.5)}$$

Eq.(A.5) tells us that the factor risk of a linear strategy is entirely due to its symmetric part, which means an antisymmetric strategy is always factor neutral. Eq.(A.2) shows instead that the expected return is impacted by both symmetric and antisymmetric parts. As long as $\Pi^a \neq 0$, an antisymmetric strategy can deliver positive expected returns without factor exposure, i.e., alphas with respect to F.

A.3 Antisymmetric Strategies

To understand the interpretation of PAPs, consider that any antisymmetric matrix $\Pi^a \in \mathbb{R}^{K\times K}$ has an even number $2K$ of eigenvalues. The nonzero eigenvalues are purely imaginary and come in complex-conjugate pairs: $i\lambda_k$ and $-i\lambda_k$. The corresponding orthonormal eigenvectors are $z_k = \frac{1}{\sqrt{2}}(x_k + iy_k)$ and the complex conjugate is $\bar{z_k} = \frac{1}{\sqrt{2}}(x_k - iy_k)$, where $x_k,\ y_k \in \mathbb{R}^N$ with $||x_k|| = ||y_k|| = 1$, $x_k'y_k = 0$ and $x_k'x_l = x_k'y_l = y_k'y_l = 0$ for all $k \neq l$,

[1] KMP provide an in-depth discussion about the economic interpretation of the latent factor. For the purpose of this paper, it is enough to notice that if $S_t = E_t[R_{t+1}]$, then F_{t+1} is the conditional tangency portfolio.

$l \leq K \leq N/2$. The corresponding eigendecomposition is given by:[2]

$$\Pi^a = \sum_{k=1}^{K} \lambda_k (x_k y_k' - y_k x_k')$$

KMP consider "rank-2 antisymmetric strategies", where $L = xy' - yx'$ for some vector $x, y \in \mathbb{R}^N$, which give rise to PAPs of the following type, as explained in the text:

$$PAP_{t+1}^k = \underbrace{S_t' x_k (y_k)' R_{t+1}}_{S_t^{x_k} R_{t+1}^{y_k}} - \underbrace{S_t' y_k (x_k)' R_{t+1}}_{S_t^{y_k} R_{t+1}^{x_k}} \tag{A.6}$$

for $k = 1, ..., N^a$ where x_k and y_k are the real and the imaginary components of the eigenvectors associated with the eigenvalues of $(\Pi^a)'$.

[2]These results are contained in Lemma 3 in KMP. Their proof is provided by the authors in the Internet Appendix to their paper.

Appendix B Cross-predictability Measures with Singular Values and Eigenvalues

From the theory (Eq.10), the k-th singular value of Π is also the expected return of the corresponding k-th PP. In practice, however, the two quantities might differ due to estimation uncertainty. Hence, as an additional estimate of the cross-predictability value based on expected returns of Section 6.1.3, I compute the percentage deviations between the first singular value of Π after performing the transformations above, and the baseline case. I therefore obtain three following alternative cross-prediction measures:

$$\text{value of Cross-prediction Signal } j := CPS_j^{\lambda} = 1 - \frac{\lambda_{-column_j}}{\lambda_{full}} \tag{B.1}$$

$$\text{value of Cross-prediction Return } j := CPR_j^{\lambda} = 1 - \frac{\lambda_{-row_j}}{\lambda_{full}} \tag{B.2}$$

$$\text{value of Cross-prediction Total } j := CPT_j^{\lambda} = 1 - \frac{\lambda_{-column_j,-row_j}}{\lambda_{full}} \tag{B.3}$$

where λ_{full} denotes the first singular value of the unrestricted Π and $\lambda_{-column_j}$, λ_{-row_j} and $\lambda_{-column_j,-row_j}$ denote the first singular value after zeroing out the j-th column, row, and column and row, respectively, except for the j-th element of the main diagonal. In sequential order CSV_j^{λ}, CRV_j^{λ} and CPV_j^{λ} measure the percentage loss in expected return from the first PP resulting from disregarding cross-prediction effects from, to, and both from and to asset j. I repeat the same procedure for the eigenvalues of Π^s and Π^a to obtain analogous measures.[3]

[3] As mentioned in the text, $E[PAP_{t+1}^k] = 2\lambda_k^a$. Hence, there is the proportionality term fades away in both CSV_j^{λ}, CRV_j^{λ} and CPV_j^{λ}.

Results are shown in Figure B.1. Let us first consider the prediction matrix Π (first panel). The patterns related to CPS^{λ} and CPT^{λ} line up almost perfectly with the $CPS^{\bar{R}}$ and $CPT^{\bar{R}}$ shown for PP in Figure 4 in the main text, with CPT_1^{λ} being a little lower than $CPT_1^{\bar{R}}$. The only difference concerns CPR^{λ}, which is now slightly positive for larger stocks, whereas it was basically zero or even slightly negative for $CPR^{\bar{R}}$. In other words, when looking at realized performance, there is virtually no value in cross-predicting stocks' returns using other assets' signals, except for small firms, even if eigenvalues suggest there should be some value. Abstracting from these small discrepancies, the similarity between results based on eigenvalues and average returns is confirmed.

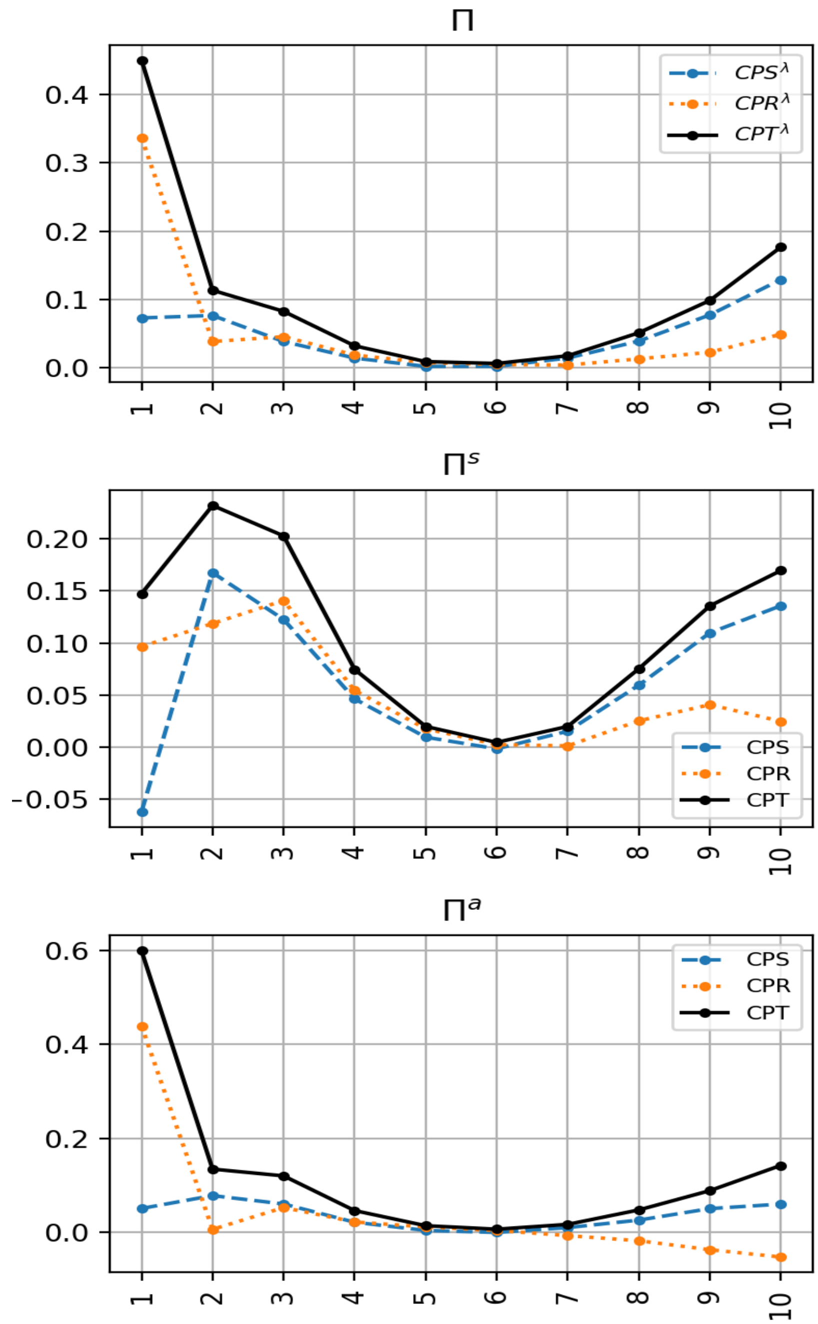

FIGURE B.1: **Size-sorted Portfolios: CSV^λ, CRV^λ and CPV^λ**

This figure shows CPS^λ, CPR^λ and CPT^λ for ten size-sorted portfolios (portfolio number on x-axis) where PPA is carried out using size as signal. λ refers to the first singular value of Π, or the first eigenvalue of Π^s and Π^a from top to bottom. The sample period is 1968 - 2019.

Appendix C Firm Signals

TABLE C.1: **Firm signals**

This table reports the 153 signals used in the paper, adapted from Table J.1 in Jensen et al. (2022).

Variable	Paper (year)
Downside beta	Ang et al. (2006)
Earnings variability	Francis et al. (2004)
Idiosyncratic volatility from the CAPM (21 days)	
Idiosyncratic volatility from the CAPM (252 days)	
Idiosyncratic volatility from the Fama-French 3-factor model	Ang et al. (2006)
Idiosyncratic volatility from the q-factor model	
Cash flow volatility	Huang (2009)
Maximum daily return	Bali et al. (2011)
Highest 5 days of return	Bali et al. (2017)
Return volatility	Ang et al. (2006)
Years 6-10 lagged returns, nonannual	Heston and Sadka (2008)
Share turnover	Datar et al. (1998)
Number of zero trades with turnover as tiedbreaker (1 month)	Liu (2006)
Number of zero trades with turnover as tiedbreaker (6 months)	Liu (2006)
Number of zero trades with turnover as tiedbreaker (12 montha)	Liu (2006)
Current price to high price over last year	George and Hwang (204)
Residual momentum t-6 to t-1	Blitz et al. (2011)
Residual momentum t-12 to t-1	Blitz et al. (2011)
Price momentum t-3 to t-1	Jegedeesh and Titman (1993)
Price momentum t-6 to t-1	Jegadeesh and Titman (1993)
Price momentum t-9 to t-1	Jegedeesh and Titman (1993)
Price momentum t-12 to t-1	Jegedeesh and Titman (1993)
Year 1-lagged return, nonannual	Heston and Sadka (2008)
Change sales minus change Inventory	Abarbanell and Bushee (1998)
Change sales minus change receivables	Abarbanell and Bushee (1998)
Change sales minus change SG&A	Abarbanell and Bushee (1998)
Change in quarterly return on assets	
Change in quarterly return on equity	
Standardized earnings surprise	Foster et al. (1984)
Change in operating cash flow to as	Bouchard et al. (2019)
Price momentum t-12 to t-7	Novy-Marx (2012)
Labor force efficiency	Abarbanell and Bushee (1998)
Standardized Revenue surprise	Jegadeesh and Livnat (2006)
Year I-lagged return, annual	Heston and Sadka (2008)
Tax expense surprise	Thomas and Zhang (2011)
Coefficient of variation for dollar trading volume	Chordia et al. (2001)
Return on net operating assets	Soliman (2008)
Profit margin	Soliman (2008)
Pitroski F-score	Piotroski (2000)
Return on equity	Haugen and Baker (1996)
Quarterly return on equity	Hou et al. (2015)
Ohlson O-score	Dichev (1998)
Operating cash flow to assets	Bouchard et al. (2019)
Operating profits-to-book equity	Fama and French (2015)
Operating profits-to-lagged book equity	
Coefficient of variation for share turnover	Chordia et al. (2001)
Capital turnover	Haugen and Baker (1996)
Cash-based operating-to-book assets	
Cash-based operating profits-to-lagged book assets	Ball et al. (2016)
Change gross margin minus change sales	Abarbanell and Bushee (1998)
Gross profits-to-assets	Novy-Marx (2013)
Gross profits-to-lagged assets	
Mispricing factor: Performance	Stambaugh and Yuan (2016)
Number of consecutive quarters with earning increases	Barth et al. (1999)
Quarterly return on assets	Balakrishnan et al. (2010)

TABLE C.1: **(continued) Firm signals**

Variable	Paper (year)
Downside beta	Ang et al. (2006)
Earnings variability	Francis et al. (2004)
Idiosyncratic volatility from the CAPM (21 days)	
Idiosyncratic volatility from the CAPM (252 days)	
Idiosyncratic volatility from the Fama-French 3-factor model	Ang et al. (2006)
Idiosyncratic volatility from the q-factor model	
Cash flow volatility	Huang (2009)
Maximum daily return	Bali et al. (2011)
Highest 5 days of return	Bali et al. (2017)
Return volatility	Ang et al. (2006)
Years 6-10 lagged returns, nonannual	Heston and Sadka (2008)
Share turnover	Datar et al. (1998)
Number of zero trades with turnover as tiedbreaker (1 month)	Liu (2006)
Number of zero trades with turnover as tiedbreaker (6 months)	Liu (2006)
Number of zero trades with turnover as tiedbreaker (12 montha)	Liu (2006)
Current price to high price over last year	George and Hwang (204)
Residual momentum t-6 to t-1	Blitz et al. (2011)
Residual momentum t-12 to t-1	Blitz et al. (2011)
Price momentum t-3 to t-1	Jegedeesh and Titman (1993)
Price momentum t-6 to t-1	Jegedeesh and Titman (1993)
Price momentum t-9 to t-1	Jegedeesh and Titman (1993)
Price momentum t-12 to t-1	Jegedeesh and Titman (1993)
Year 1-lagged return, nonannual	Heston and Sadka (2008)
Change sales minus change Inventory	Abarbanell and Bushee (1998)
Change sales minus change receivables	Abarbanell and Bushee (1998)
Change sales minus change SG&A	Abarbanell and Bushee (1998)
Change in quarterly return on assets	
Change in quarterly return on equity	
Standardized earnings surprise	Foster et al. (1984)
Change in operating cash flow to as	Bouchard et al. (2019)
Price momentum t-12 to t-7	Novy-Marx (2012)
Labor force efficiency	Abarbanell and Bushee (1998)
Standardized Revenue surprise	Jegadeesh and Livnat (2006)
Year I-lagged return, annual	Heston and Sadka (2008)
Tax expense surprise	Thomas and Zhang (2011)
Coefficient of variation for dollar trading volume	Chordia et al. (2001)
Return on net operating assets	Soliman (2008)
Profit margin	Soliman (2008)
Pitroski F-score	Piotroski (2000)
Return on equity	Haugen and Baker (1996)
Quarterly return on equity	Hou et al. (2015)
Ohlson O-score	Dichev (1998)
Operating cash flow to assets	Bouchard et al. (2019)
Operating profits-to-book equity	Fama and French (2015)
Operating profits-to-lagged book equity	
Coefficient of variation for share turnover	Chordia et al. (2001)
Capital turnover	Haugen and Baker (1996)
Cash-based operating-to-book assets	
Cash-based operating profits-to-lagged book assets	Ball et al. (2016)
Change gross margin minus change sales	Abarbanell and Bushee (1998)
Gross profits-to-assets	Novy-Marx (2013)
Gross profits-to-lagged assets	
Mispricing factor: Performance	Stambaugh and Yuan (2016)
Number of consecutive quarters with earning increases	Barth et al. (1999)
Quarterly return on assets	Balakrishnan et al. (2010)

Table C.1: **(continued) Firm signals**

Variable	Paper (year)
Operating profits-to-book assets	
Operating profits-to-lagged book assets	Ball et al. (2016)
Operating leverage	Novy-Marx (2011)
Quality minus Junk: Composite	Assness et al. (2018)
Quality minus Junk: Growth	Assness et al. (2018)
Quality minus Junk: Profitability	Assness et al. (2018)
Quality minus Junk: Safety	Assness et al. (2018)
Assets turnover	Soliman (2008)
Market correlation	Assness et al. (2020)
Coskewness	Harvey and Siddique (2000)
Net debt issuance	Bradshaw et al. (2006)
Kaplan-Zingales index	Lamont et al. (2001)
Change in long-term investments	Richardson et al. (2005)
Taxable income-to-book income	Lev and Nissim (2004)
Years 2-5 lagged returns, annual	Heston and Sadka (2008)
Years 6-10 lagged returns, annual	Heston and Sadka (2008)
Years 11-15 lagged returns, annual	Heston and Sadka (2008)
Years 11-15 lagged returns, nonannual	Heston and Sadka (2008)
Years 16-20 lagged returns, annual	Heston and Sadka (2008)
Change in short-term investments	Richardson et al. (2005)
Amihud Measure	Amihud (2002)
Dollar trading volume	Brennan et al. (1998)
Market Equity	Banz (1981)
Price per share	Miller and Scholes (1982)
R&D-to-market	Chan et al. (2001)
Idiosyncratic skewness from the CAPM	
Idiosyncratic skewness from the Fama-French 3-factor model	Bali et al. (2016)
Idiosyncratic skewness from the q-factor model	
Short-term reversal	Jegadeesh (1990)
Highest 5 days of return scaled by volatility	Assness et al. (2020)
Total skewness	Bali et al. (2016)
Assets-to-market	Fama and French (1992)
Book-to-market equity	Rosenberg et al. (1985)
Book-to-market enterprise value	Penman et al. (2007)
Net stock issues	Pontiff and Woodgate (2008)
Debt-to-market	Bhandari (1988)
Dividend yield	Litzenberger and Ramaswamy (1979)
Ebitda-to-market enterprise value	Loughran and Wellman (2011)
Equity duration	Dechow et al. (2004)
Net equity issuance	Bradshaw et al. (2006)
Equity net payout	Daniel and Titman (2006)
Net payout yield	Boudoukh et al. (2007)
Payout yield	Boudoukh et al. (2007)
Free cac;h flow-to-price	Lakonishok et al. (1994)
Intrinsic value-to-market	Frankel and Lee (1998)
Net total issuance	Bradshaw et al. (2006)
Earnings-to-price	Basu (1983)
Operating cac;h flow-to-market	Desai et al. (2004)
Sales-to-market	Barbee et al. (1996)

Appendix D Double-sorted Portfolios

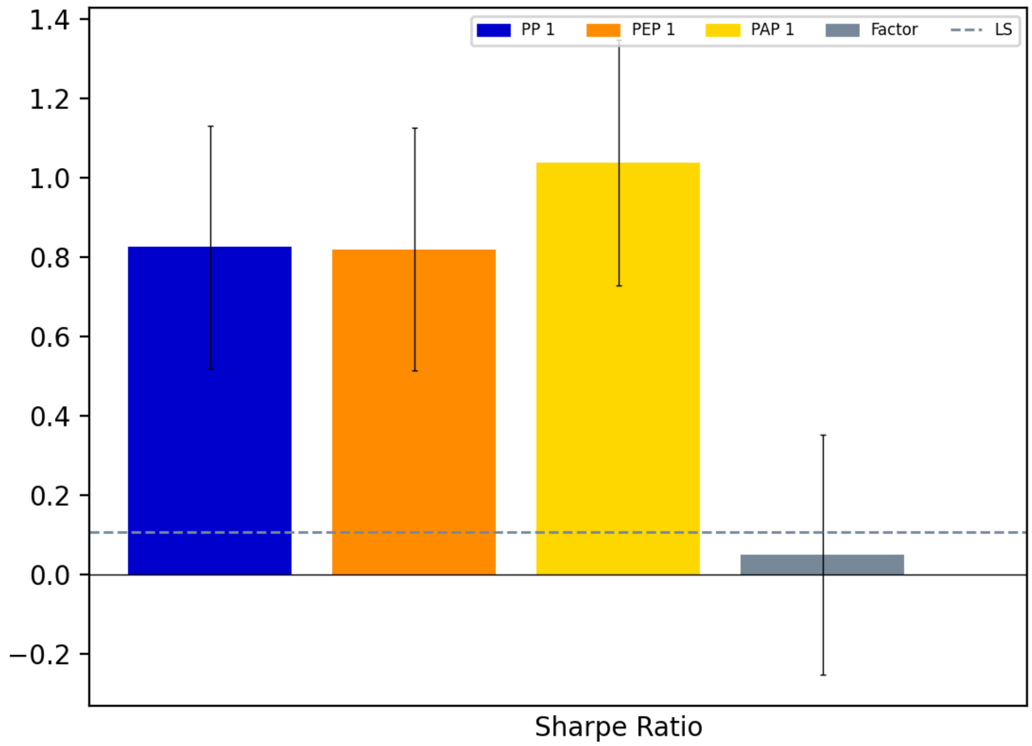

Figure D.1: **Performance ratios, size-and-book-to-market-sorted portfolios**

This figure shows average out-of-sample annualized SRs over 153 signals for the first PP, PEP and PAP, along with ±2 standard error bands around each estimate for 25 size-and-book-to-market-sorted portfolios. PPA is carried out using size as signal. Each forecast is made on an out-of-sample basis using a rolling training window of 120 months. The sample period is 1968 - 2019.

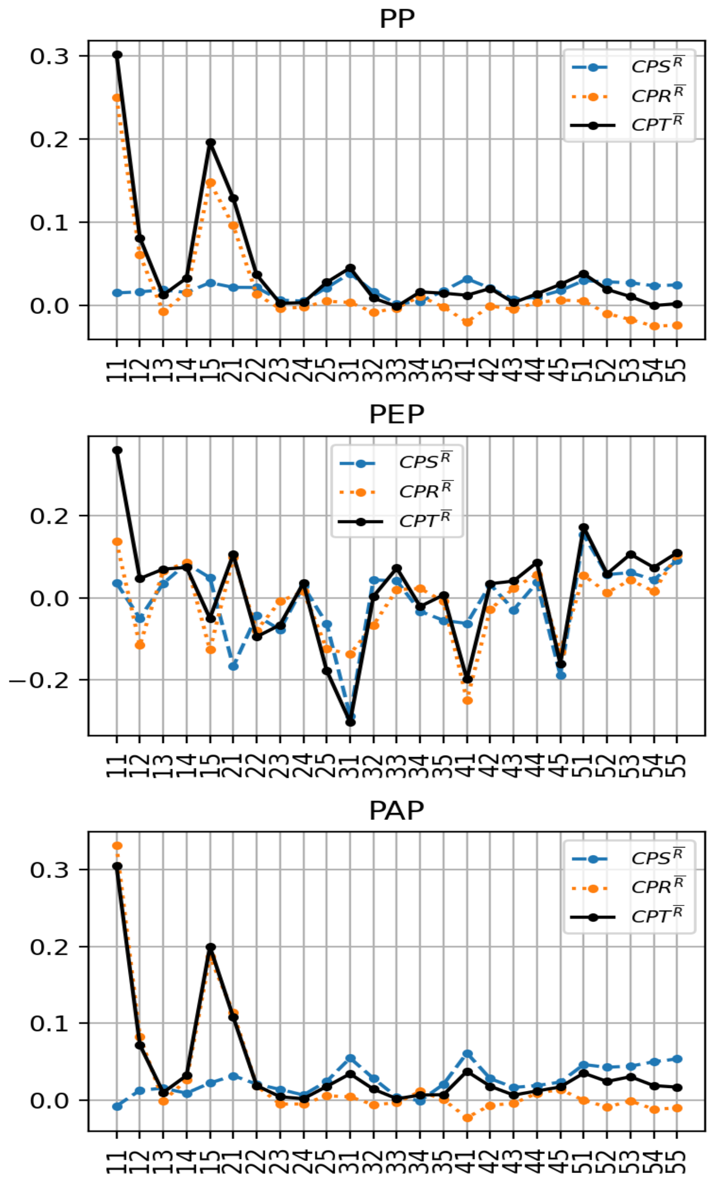

FIGURE D.2: **Size-and-book-to-market-sorted Portfolios: $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ across all signals**

This figure shows the average $CPS^{\bar{R}}$, $CPR^{\bar{R}}$ and $CPT^{\bar{R}}$ for twenty-five size-and-book-to-market-sorted portfolios (portfolio number on x-axis) over 153 signals, where PPA is carried out using one signal at a time. The first digit identifies the portfolio number along the size dimension and the second one along the book-to-market dimension. Realized returns refer to a trading strategy based on the first PP, PEP, and PAP, from top to bottom. The sample period is 1968 - 2019.

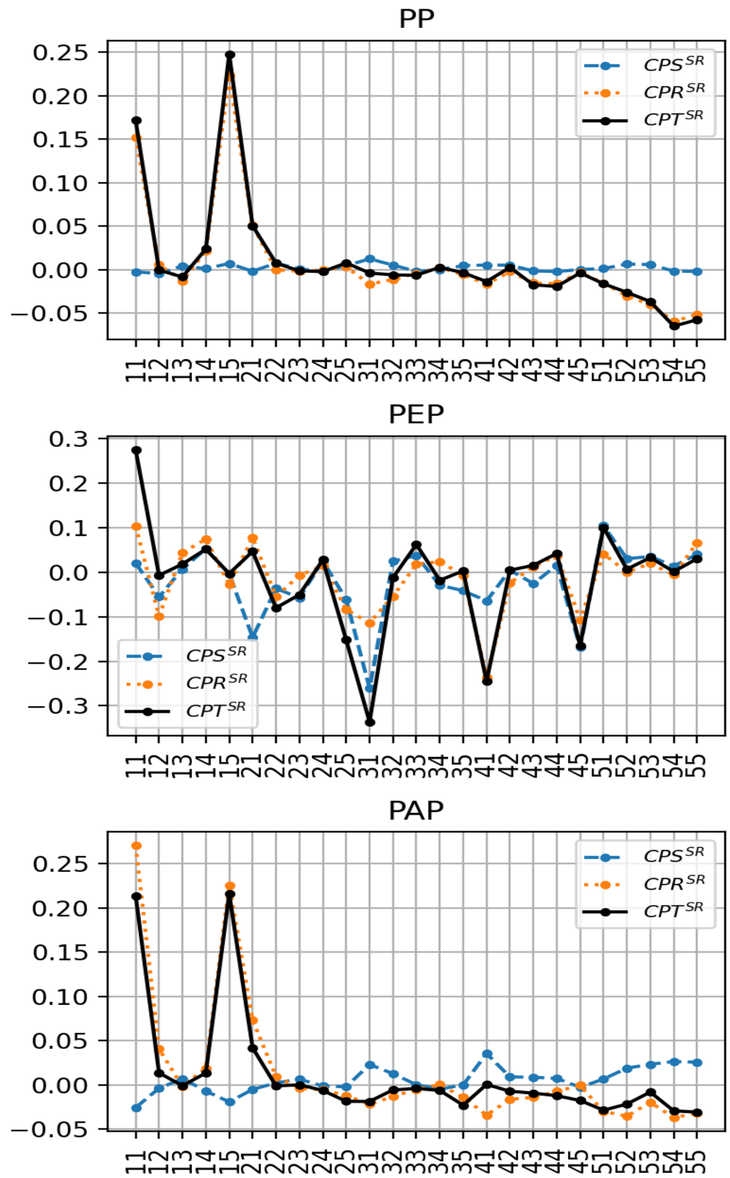

FIGURE D.3: **Size-and-book-to-market-sorted Portfolios: CPS^{SR}, CPR^{SR} and CPT^{SR} across all signals**

This figure shows the average CPS^{SR}, CPR^{SR} and CPT^{SR} for twenty-five size-and-book-to-market-sorted portfolios (portfolio number on x-axis) over 153 signals, where PPA is carried out using one signal at a time. The first digit identifies the portfolio number along the size dimension and the second one along the book-to-market dimension. SR refers to a trading strategy based on the first PP, PEP and PAP, from top to bottom. The sample period is 1968 - 2019.

Table D.1: **Cross-predictability: Fama-MacBeth Regressions (Average Returns), Size-and-book-to-market-sorted portfolio**

This table reports the results of Fama-MacBeth (Fama and MacBeth, 1973) cross-sectional regressions according to Eq.(21). The first, second and third panel refer to $CPT^{\bar{R}}$, $CPS^{\bar{R}}$ and $CPR^{\bar{R}}$ based on rolling average returns for twety-five size-and-book-to-market-sorted portfolios as dependent variable, respectively. MIS_t is the $MGMT$ mispricing measure from Stambaugh and Yuan (2017) and X_t is a vector of controls which includes portfolio-level book-to-market ratio (BM), momentum (MOM) and operating profitability (OP) in column (1), together with idiosyncratic volatility ($IVOL$) in column (2) and idiosyncratic volatility and Size in column (3). Both dependent and independent variables are normalized by their standard deviation. t-statistics based on Newey-West standard errors (Newey and West, 1987) with 4 lags (chosen following Greene (2003)) are in brackets. Bold numbers indicate significant statistics based on conventional significance levels. The last column of each panel reports the average cross-sectional R^2. Accounting for the rolling estimation, the available sample period is 1989 - 2019.

$CPT^{\bar{R}}$	(1)	(2)	(3)
MIS	**-0.03 (-3.93)**	**-0.03 (-3.9)**	**-0.03 (-4.32)**
$IVOL$		**0.09 (2.07)**	**0.12 (2.76)**
$Size$			**-1.17 (-2.93)**
BM	**-0.26 (-7.92)**	**-0.31 (-7.36)**	**-0.34 (-8.7)**
MOM	0.06 (0.79)	0.05 (0.66)	0.11 (1.28)
OP	**-0.02 (-1.94)**	-0.02 (-1.43)	-0.01 (-0.86)
R^2	0.05	0.04	0.08
$CPS^{\bar{R}}$	(1)	(2)	(3)
MIS	-0.03 (-1.62)	-0.02 (-1.43)	**-0.06 (-3.32)**
$IVOL$		**-0.74 (-10.22)**	**-0.35 (-5.27)**
$Size$			**3.3 (9.96)**
BM	**-0.42 (-7.62)**	**-0.21 (-2.87)**	**-0.29 (-4.22)**
MOM	**1.2 (8.37)**	**1.19 (7.98)**	**1.09 (6.77)**
OP	**0.15 (6.72)**	**0.16 (6.66)**	**0.19 (8.2)**
R^2	0.08	0.11	0.14
$CPR^{\bar{R}}$	(1)	(2)	(3)
MIS	0.04 (0.57)	0 (0.26)	0 (0.38)
$IVOL$		**0.23 (4.88)**	**0.24 (4.72)**
$Size$			**-1.61 (-3.72)**
BM	**-1.84 (-4.51)**	**-0.36 (-7.23)**	**-0.37 (-8.42)**
MOM	-0.27 (-0.35)	-0.12 (-1.28)	-0.05 (-0.54)
OP	**-0.89 (-8.08)**	**-0.1 (-7.29)**	**-0.09 (-7.75)**
R^2	0.1	0.14	0.17

Table D.2: **Cross-predictability: Fama-MacBeth Regressions (SRs), Size-and-book-to-market-sorted portfolios**

This table reports the results of Fama-MacBeth (Fama and MacBeth, 1973) cross-sectional regressions according to Eq.(21). The first, second and third panel refer to CPT^{SR}, CPS^{SR} and CPR^{SR} based on rolling average SRs for twenty-five size-and-book-to-market-sorted portfolios as dependent variable, respectively. MIS_t is the $MGMT$ mispricing measure from Stambaugh and Yuan (2017) and X_t is a vector of controls which includes portfolio-level book-to-market ratio (BM), momentum (MOM) and operating profitability (OP) in column (1), together with idiosyncratic volatility ($IVOL$) in column (2) and idiosyncratic volatility and Size in column (3). Both dependent and independent variables are normalized by their standard deviation. t-statistics based on Newey-West standard errors (Newey and West, 1987) with 4 lags (chosen following Greene (2003)) are in brackets. Bold numbers indicate significant statistics based on conventional significance levels. The last column of each panel reports the average cross-sectional R^2. Accounting for the rolling estimation, the available sample period is 1989 - 2019.

CPT^{SR}	(1)	(2)	(3)
MIS	0.01 (1.16)	0.01 (0.98)	0 (-0.45)
$IVOL$		**0.18 (3.58)**	**0.21 (3.94)**
$Size$			**-1.6 (-3.02)**
BM	**-0.32 (-12.47)**	**-0.42 (-9.9)**	**-0.46 (-12)**
MOM	-0.05 (-0.64)	-0.06 (-0.84)	0.01 (0.1)
OP	**-0.16 (-12.82)**	**-0.14 (-12.05)**	**-0.12 (-11.75)**
R^2	0.24	0.25	0.3
CPS^{SR}	(1)	(2)	(3)
MIS	**-0.07 (-5.62)**	**-0.07 (-5.74)**	**-0.08 (-5.61)**
$IVOL$		**-0.14 (-2.68)**	0.02 (0.29)
$Size$			**0.94 (4.71)**
BM	**0.17 (3.77)**	**0.22 (3.71)**	**0.16 (2.36)**
MOM	0.09 (1.09)	0.06 (0.73)	0.1 (1.05)
OP	**0.13 (10.27)**	**0.13 (9.63)**	**0.14 (10.29)**
R^2	0.16	0.19	0.2
CPR^{SR}	(1)	(2)	(3)
MIS	0.01 (1.51)	0.01 (1.42)	0 (0.12)
$IVOL$		**0.28 (5.32)**	**0.31 (5.61)**
$Size$			**-1.61 (-2.93)**
BM	**-0.33 (-12.16)**	**-0.45 (-9.7)**	**-0.49 (-11.82)**
MOM	**-0.15 (-1.95)**	**-0.15 (-2.01)**	-0.1 (-1.25)
OP	**-0.16 (-12.16)**	**-0.15 (-10.92)**	**-0.13 (-10.86)**
R^2	0.26	0.29	0.33

Recent Issues

No. 423	Tobias Berg, Florian Heider	Leverage and Risk-Taking in a Dynamic Model
No. 422	Julian Detemple	Thoughts about the Dictator and Trust Game
No. 421	Claes Bäckman, Patrick Moran, Peter van Santen	Mortgage Design, Repayment Schedules, and Household Borrowing
No. 420	Konstantin Bräuer, Andreas Hackethal, Tobin Hanspal, Samuel M. Hartzmark	Educating Investors about Dividends
No. 419	Andreas Hackethal, Philip Schnorpfeil, Michael Weber	Inflation and Trading
No. 418	Ioannis Asimakopoulos, Tobias Tröger	Reform of the CMDI Framework – Driving Off With the Breaks On
No. 417	Thiago Fauvrelle, Max Riedel, Mathias Skrutkowski	Collateral Pledgeability and Asset Manager Portfolio Choices During Redemption Waves
No. 416	Olga Balakina, Claes Bäckman, Anastasiia Parakhoniak	Beyond Connectivity: Stock Market Participation in a Network
No. 415	Monica Billio, Aoife Fitzpatrick, Carmelo Latino, Loriana Pelizzon	Unpacking the ESG Ratings: Does One Size Fit All?
No. 414	Peter Andre, Teodora Boneva, Felix Chopra, Armin Falk	Misperceived Social Norms and Willingness to Act Against Climate Change
No. 413	Errikos Melissinos	Real Term Premia in Consumption-Based Models
No. 412	Andreas Hackethal, Thomas Pauls	Conform to the Norm. Peer Information and Sustainable Investments
No. 411	Peter Andre, Marc Kaufmann, Botond Köszegi	Understanding Markets with Socially Responsible Consumers
No. 410	Carmelo Latino	Surfing the Green Wave: What's in a "Green" Name Change?
No. 409	Kevin Bauer, Yan Chen, Florian Hett, Michael Kosfeld	Group Identity and Belief Formation: A Decomposition of Political Polarization