Assaying Anomalies*

Robert Novy-Marx† Mihail Velikov‡

February 13, 2024

Abstract

We propose a protocol for testing potential cross-sectional predictors of equity returns, and describe turn-key tools for democratizing the implementation of protocol with little more effort than pushing a button. Our free-to-use web application automatically generates an online appendix with text, tables, and figures, analyzing the performance of a candidate cross-sectional return predictor. The tests in our protocol go far beyond the direct inferences available from standard linear factor models, identifying issues that commonly arise testing equity strategies, paying particular attention to arbitrage limits that can make a strategy look good on paper even when if cannot be profitably traded in practice. It also identifies similar anomalies and places the proposed predictor in the context of the now extensive "factor zoo."

JEL classification: G11, G12, G14
Keywords: Anomalies, Performance Evaluation, Trading Costs, Factor Models

---

*Novy-Marx provides consulting services to Dimensional Fund Advisors, an investment firm headquartered in Austin, Texas with strong ties to the academic community. The thoughts and opinions expressed in this article are those of the authors alone, and no other person or institution has any control over its content.

†Simon Business School and NBER, University of Rochester, 500 Joseph C. Wilson Blvd., Box 270100, Rochester, NY 14627. Email: robert.novy-marx@simon.rochester.edu.

‡Smeal College of Business, Pennsylvania State University, 366 Business Building, University Park, PA 16802. Email: velikov@psu.edu.

> **assay** (as•say) *Verb* • To analyze (something, such as an ore) for one or more specific components. • To judge the worth of.
>
> Merriam-Webster Dictionary

1 Introduction

This paper describes a protocol, and easily-accessible, easily-implementable tools, for dissecting and understanding newly proposed cross-sectional equity return predictors. While simple tools cannot be completely exhaustive, they should identify the most important issues that arise in common tests of asset pricing strategies, while going far beyond the direct inferences available from the simple tests commonly employed using standard linear factor models. The tools described here automatically generate a complete paper, providing a transparent analysis along the lines of a thorough referee report, with little more effort than pushing a button.

After its introduction, the Fama and French (1993) three-factor model quickly became the dominant asset pricing model. It was swiftly adopted not only because it allowed researchers to quickly diagnose hidden tilts to the most common empirical phenomena of the day, but also because of the simplicity with which it could be operationalized.

The model's dominance has, however, created its own set of problems. Broad professional agreement on a standard model created a widely accepted, simple definition of an "anomaly," as any strategy that has abnormally high returns relative to the model. Anomalous returns became an almost necessary condition for publishing asset pricing papers. This created tremendous incentives for researchers to find these anomalies, often without any grounding in underlying economic theory, at a time when the profession was experiencing explosive growth, and technological innovation had lowered the bar on testing. The result is our current "factor zoo" (Cochrane, 2011), with hundreds of documented cross-section predictors of stock returns (Harvey et al., 2016).

The large, growing number of known predictors has itself also made it increasingly challenging to evaluate the contribution and robustness of newly proposed candidates for the factor zoo. The simple diagnostic tests allowed by the original Fama and French (1993) model and its extensions are now nowhere near as informative as they were upon its introduction. With the large number of known anomalies it is impossible for a three-factor model, or even a five- or six-factor model, to uncover all the potential tilts to known phenomena.

Perhaps more importantly, the profession has uncovered many techniques that appear to increase the magnitude of anomalous returns, but do so in ways that are economically uninteresting and can be difficult to detect. These are mostly driven by implementation issues, which limit the forces of arbitrage. The most obvious of these involve over-weighting small cap stocks. Anomalies almost always appear stronger, often far stronger, among small stocks, at least when one ignores transaction costs. While these strategies are more expensive to trade, and consequently don't represent a more attractive trading opportunity or attract significant arbitrage capital (e.g., Novy-Marx and Velikov, 2016), they generate large gross alphas, allowing researchers to report high statistical significance. Portfolio weighting schemes that overweight smaller stocks in non-obvious ways consequently gain popularity, because they contribute to impressive paper performance (see, e.g., Velikov and Novy-Marx, 2022). Similarly, more frequent trading can improve the paper performance of anomalies, but also entails significant, largely ignored costs. The tools presented here explicitly account for implementation costs and should consequently be of interest to practitioners as well as academics.

Increased computing power has also enabled more sophisticated methods for summarizing and evaluating data. These machine learning techniques are especially important in the presence of the factor zoo, as they offer tools for imposing sparsity. They can help researchers select only an important subset of many potential factors under consideration to use in their analysis, or reduce dimensionality by coming up with particularly important combinations of factors. These techniques, while becoming more popular in finance, are still not a part of most researchers toolbox. While not our main focus, our project incorporates tests employing some of the machine learning techniques into our analysis.

The tools that run all these tests and automatically generate a report will be soon available in two forms: a free web application, and a public github repository.[1] The web application allows users to test the robustness of a new predictive signal by uploading a .csv file with three columns: firm identifier, date, and signal. The application then generates a self-contained report testing the new anomaly, and emails the submitter latex files and .pdf documents for this report. This referee-style report includes extensive diagnostic and robustness results, as well as an estimation of a taxonomic rank that places the proposed anomaly in context in the factor zoo, described in our protocol. The advantage of this modality is its ease of use. It is accessible to everyone, and does not require any coding skills or the use of a particular platform.

The public github repository contains an extensive library of MATLAB code that implements the tests from scratch with just a couple of mouse clicks. While these tools are based on a specific platform, and using them requires a little more skill and subscriptions to the usual data vendors, this modality has several advantages.[2]

[1] A fully operational preliminary version of the web application , tutorials on the github repository, and additional samples of the automatically generated reports are available at http://assayinganomalies.com/. The github repository is available at https://github.com/velikov-mihail/AssayingAnomalies.

[2] We are currently investigating having all these tools also translated into Python.

These tools are more flexible. They can be modified and adapted by individual users for their specific needs. This also allows for the tools' open-source evolution over time. Moreover, the github repository contains a far broader set of tools, offering functionality that goes far beyond the actual testing protocol proposed here. It includes tools for accessing and downloading data from common sources, organizing this data, and running common tests in the asset pricing literature. These are all well-documented, and designed so that their basic functionality requires minimal coding skill. This dramatically lowers the bar for researchers wanting to start serious empirical work, offering an easy on-ramp for those beginning their careers. Finally, the tools interface with related public github repositories, giving access to a growing library of replications of important (and not so important) papers in the literature, and code for running empirically driven finance classes.

Our goal is to provide the option for any authors proposing new anomalies to freely implement our protocol with minimal effort. The potential impact of the project on the academic literature is extensive. The protocol provides a common, easily-accessible framework for the basic testing for new factors. This removes degrees of freedom that authors have when testing proposed new anomalies, thereby mitigating the overfitting concerns that have become increasingly pernicious for the profession. It makes it easy for reviewers to request, and authors to provide, a set of standard robustness checks for an online appendix. The project also has significant practical relevance due to its emphasis on accounting for implementation frictions, and can help bridge the gap between academic research on new factors and their application in industry.

2 Walk-through of the anomaly testing protocol

The following gives a brief overview of the actual tests performed and exhibits produced by the code when it automatically generates its report. Appendix A provides more details, presenting an actual example of an automatically generated report using input data for one of our own published anomalies, the monetary policy exposure index (MPE) from Ozdagli and Velikov (2020). An overview of this report is provided below.

2.1 Section 1: Introduction

Section 1 briefly describes the report and how it is generated, referencing this paper's protocol. It states the specific version of the publicly available code that was used to produce the report.

2.2 Section 2: Signal diagnostics

Section 2 provides signal diagnostics. Figure 1 plots descriptive statistics for the proposed predictor (Panel A) and its coverage over time both as a fraction of total firms and total market capitalization (Panel B). The plot helps identify any obvious outliers and if there are any periods with poor data coverage.

2.3 Section 3: Does the signal predict returns?

Section 3 checks whether the signal reliably predicts cross-sectional differences in average returns. Table 1 reports time-series regression results employing the value-weighted returns to portfolios constructed from a quintile sort using NYSE breakpoints on the candidate predictor (MPE for this example). Univariate sorts like these are the main technique in the anomaly literature to test whether a signal predicts returns in the cross-section of equities. A version of this table is what most anomaly papers report, though they vary in the specific portfolio construction. Our choice of value-weighting and NYSE breakpoints is conservative, as anomalies are usually strongest among micro-cap stocks and thus generally look stronger when implemented using equal-weighted portfolio returns or name breakpoints (Fama and French, 2008). Our default choice of value-weighting and NYSE breaks provides results that are closer to what an actual investor might be able to achieve in practice.

Table 2 reports results for various alternative construction methodologies. It varies the number of portfolios (five or ten), the type of portfolio breakpoints employed (NYSE, name, or capitalization), and the weighting of individual stocks within each portfolio (value- or equal-weighting). Panel B of Table 2 considers the impact of accounting for transaction costs. The trading cost calculation follows Detzel et al. (2022). The net-of-costs return on anomaly f in month t is given by:

$$f_t^{net} = f_t^{gross} - TC_{Long,t} - TC_{Short,t}$$

where

$$TC_{j,t} = \sum_{i \in I_{j,t}} |w_{i,t} - \tilde{w}_{i,t-1}| \cdot c_{i,t}$$

for $j \in \{Long, Short\}$ and $I_{j,t}$ indexes the stocks in portfolio j at time t; $c_{i,t}$ is the one-way trading cost of stock i at time t, measured as the high-frequency combination effective half-spreads from Chen and Velikov (2022); $w_{i,t}$ is the weight of stock i in its portfolio at time t after rebalancing and $\tilde{w}_{i,t-1} = \frac{w_{i,t-1}(1+r_{it})}{\sum_{k \in I_{j,t}} w_{k,t-1}(1+r_{kt})}$ is the weight of the stock in the portfolio before rebalancing.

Table 2, Panel B also reports the Novy-Marx and Velikov (2016) generalized alphas that account for trading costs. It reports these generalized alphas relative to five models: the CAPM, the Fama and French three- and five-factor models, and these three- and five-factor models augmented with the momentum factor UMD. The alphas are estimated as

$$w_{y,MVE_{\{X,y\}}}^{-1} MVE_{\{X,y\}} \;=\; \alpha^* + \beta^* \cdot MVE_{\{X\}} + \epsilon^*,$$

where $MVE_{\{X\}}$ denotes the ex-post mean-variance efficient portfolio of the assets X, where X are the factors in the model, and $w_{y,MVE_{\{X,y\}}}$ denotes the weight on asset y (the candidate factor) in $MVE_{\{X,y\}}$. Following Novy-Marx and Velikov (2016), α^* is defined to equal 0 when $w_{y,MVE_{\{X,y\}}} = 0$.

Table 3 explicitly accounts for the role of firm size in the strength of the candidate anomaly's performance. It does so by constructing strategies based on the candidate cross-sectional returns predictor within NYSE size quintiles. The table reports average portfolio returns, average number stocks, and average firm size, for twenty five portfolios constructed from a conditional double sort on size and and the proposed signal. It also reports the average returns and alphas for long/short trading strategies based on the signal within each size quintile.

2.4 Section 4: Signal performance relative to the factor zoo

Section 4 considers the strategy's performance in the context of the factor zoo. It does so by comparing the proposed factor's performance to that of 207 anomalies from the literature satisfying our criteria for inclusion in the testing protocol.[3]

Figure 2 plots histograms of gross and net Sharpe ratios for 207 known anomalies and places the candidate factor in these distributions (Panel A and B, respectively). To keep performances comparable, SR for anomalies in the factor zoo are calculated over the sample for which the candidate return predictor is provided.

Figure 3 plots the growth of a $1 invested in each of the the 207 known anomalies, and compares those with the growth of a $1 invested in the test signal strategy (red

[3]The anomalies come from the March, 2022 release of the Chen and Zimmermann (2022) open source asset pricing dataset.

lines), again on both a gross and net basis (top and bottom panels, respectively).[4]

Figure 4 shows how the candidate strategy performs relative to known anomalies in expanding the investment frontier spanned by common factor models. It plots the entire distribution, from lowest to highest, of the gross alphas (left panel) and the net generalized alphas (right hand panel) for each anomaly in the factor zoo relative to each of the five models used in Table 1 (CAPM and Fama-French three-, four-, five-, and six-factor models). It then places gross and net generalized alphas of the candidate strategy relative to each of these models in these distributions.

2.5 Section 5: Does the signal add relative to related anomalies?

Even if a candidate strategy has strong performance relative to most of the factors in the zoo, it may still not add significantly to the factor zoo. For example, a slight variation on one of the strongest know anomalies will itself have strong performance, but will not be a significant addition to the zoo already containing the strategy on which it is a variation. Section 5 accounts for this, by checking if the test signal adds information beyond that provided to the most closely related known anomalies.

Figures 5 and 6 show how closely related the candidate strategy is to members of the factor zoo. Figure 5 plots a name histogram of the panel correlations of the test signal with the anomaly signals from the factor zoo. Figure 6 shows an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

Figure 7 shows how much the candidate signal adds relative to each individual member in the factor zoo. It plots histograms of t-statistics for predictability tests, which test the power of the test signal controlling for other individual known anomaly

[4] If an anomaly in the factor zoo starts later than the candidate strategy, then for that factor we assume that the dollar is invested in the candidate strategy up to the date the factor becomes available.

signals. Panel A reports t-statistics on the loading on the test signal, $t(\beta_S)$ from Fama-MacBeth regressions of the form:

$$r_{i,t} = \alpha + \beta_S S_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$$

where X stands for one of the anomaly signals at a time, and S stands for the test signal.

Panel B plots t-statistics on α from spanning tests of the form:

$$r_{S,t} = \alpha + \beta r_{X,t} + \epsilon_t$$

where $r_{X,t}$ stands for the returns to one of the anomaly trading strategies at a time, and $r_{S,t}$ stands for the returns to the test signal trading strategy. The strategies employed in the spanning tests are constructed using quintile sorts, value-weighting, and NYSE breakpoints.

Panel C plots t-statistics on the average returns to strategies constructed by conditional double sorts. In each month, we sort stocks into quintiles based one of the anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on the test signal. Stocks are finally grouped into five test-signal-portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted trading strategies of the test signal conditioned on each of the anomalies.

Tables 4 and 5 control for the six most-closely related anomalies. To find the most closely related anomalies, we rank all anomalies based on:

$$\text{rank}(|\rho_{i,s}|) + rank(R^2_{r^i_t = \alpha + \beta r^i_s + \epsilon}),$$

where $\rho_{i,s}$ is the panel correlation of the underlying signal for anomaly i and the test signal s, and $R^2_{r^i_t=\alpha+\beta r^i_s+\epsilon}$ is R^2 from the spanning test of regressing the returns to the testing strategy exploiting anomaly i on the test signal s.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on the test signal controlling for the six most closely-related anomalies, both individually and jointly. Table 5 reports spanning tests results from time-series regressions of the returns to the test signal trading strategy onto the returns of trading strategies exploiting the six most closely-related anomalies and the six Fama-French factors.

2.6 Section 6: Does the signal add relative to the whole zoo?

Section 6 quantifies the extent to which the test signal increases the investment frontier beyond that spanned by the entire factor zoo.

Figure 8 plots the growth of a \$1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). We combine signals using a linear model of expected returns:

$$\mathbb{E}_t(r_{i,t+1}) = \beta_0 + \sum_{j=1}^{J} \beta_j x_{i,j,t},$$

where $r_{i,t+1}$ is the gross return of stock i in month $t+1$, J is the total number of predictors, β_j is the slope coefficient on predictor j, and $x_{i,j,t}$ is the standardized jth anomaly characteristic for stock i in month t.[5]

The figure shows results using six different methods for combining anomalies. The methods used are average rank (i.e., $\hat{\beta}_j = \frac{1}{J}$), weighted-average rank (i.e., $\hat{\beta}_j \propto \bar{r}^j$), Fama-MacBeth regression following Lewellen (2015), Partial Least Squares (PLS) filter following Light et al. (2017), Instrumented Principal Component Analysis (IPCA) following Kelly et al. (2019), and the Least Absolute Shrinkage and Selection Oper-

[5]For these combination signals, we filter the 207 anomalies and require for each anomaly the average month to have at least 40% of the cross-sectional observations available for market capitalization on CRSP in the period for which the candidate signal is available.

ator (LASSO) as implemented in Chen and Velikov (2022).

The figure compares the performance of combinations made using the broad cross-section of known anomalies, and the extent to which performance is improved by also including the proposed candidate.

3 Discussion and caveats

First, as mentioned previously, the protocol is not meant to be completely exhaustive. The tools provide a thorough, transparent analysis, going far beyond the tests commonly employed using standard linear factor models. They cannot, however, account for everything, and that is not their purpose. They are meant to effortlessly identify the most important issues that arise in common tests of asset pricing strategies. They are also specifically designed to identify strategies constructed to exploit market frictions that limit arbitrage in ways that strengthen paper performance without improving the opportunities available to actual investors.

Discrete signals also present a challenge, especially when the set of possible values is small. Many of our tests rely on non-parametric methods involving assigning stocks to portfolios on the basis of some signal. When many firms share the same signal, then some firms with the same signal must be assigned to different portfolios or the resulting portfolios will be unbalanced. In our context, where the tests are designed to run independently without requiring human judgement specific to the signal being tested, thin portfolio present a real risk. When multiple firms can naturally be assigned to two different portfolios, we consequently let nature chose which firms are assigned to each in a manner that ensures a similar degree of portfolio diversification. That is, we have some random assignment among firms with identical signals. While there is nothing inherently wrong with this procedure, it is somewhat arbitrary, and complicates the interpretation of results involving discrete signals.

Finally, even when the protocol uncovers serious inconsistencies across differently constructed strategies formed on the basis of a candidate predictor, the underlying signal may still be interesting. Several of the tests are designed to identify difficulties that may arise exploiting the strategy in practice due to market frictions. Results that reveal significant differences in performance across construction methods point to significant implementation issues related to limits to arbitrage. While this does suggest the strategy may be of limited interest to practitioners as an investment opportunity, the existence and nature of the limits to arbitrage that impact strategies based on the signal may themselves be highly interesting.

4 Conclusion

This paper describes a protocol for testing potential cross-sectional equity return predictors. This protocol goes far beyond the simple tests commonly employed using standard linear factor models, and identifies the most important issues that arise testing asset pricing strategies. It also describes turn-key tools for implementing this protocol, which produce a thorough, transparent analysis, along the lines of a referee report, with little more effort than pushing a button. These are part of broader package of freely available tools offering functionality that goes far beyond the testing protocol proposed here. These are well-documented and require minimal coding skill, dramatically lowering the bar for researchers wanting to start serious empirical work and offering an easy on-ramp for those beginning their careers.

References

Chen, A. and Velikov, M. (2022). Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis*, Forthcoming.

Chen, A. Y. and Zimmermann, T. (2022). Open source cross-sectional asset pricing. *Critical Finance Review*, 27(2):207–264.

Cochrane, J. (2011). Discount rate. *Journal of Finance*, 66(4):1047–1108.

Detzel, A., Novy-Marx, R., and Velikov, M. (2022). Model comparison with transaction costs. *Journal of Finance, Forthcoming*.

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

Fama, E. F. and French, K. R. (2008). Dissecting anomalies. *Journal of Finance*, 63(4):1653–1678.

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

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

Lewellen, J. (2015). The cross-section of expected returns. *Critical Finance Review*, 4(1):1–44.

Light, N., Maslov, D., and Rytchkov, O. (2017). Aggregation of information about the cross section of stock returns: A latent variable approach. *Review of Financial Studies*, 30:1339–1381.

Novy-Marx, R. and Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies*, 29(1):104–147.

Ozdagli, A. and Velikov, M. (2020). Show me the money: The monetary policy risk premium. *Journal of Financial Economics*, 135(2):320–339.

Velikov, M. and Novy-Marx, R. (2022). Betting against betting against beta. *Journal of Financial Economics*, 143(1):80–106.

Appendix A Output example

This Appendix provides the actual output from the tools that implement our protocol. These take a flat .csv data file with three columns—firm identifier, date, and signal—and generate a .tex file for a referee report or internet appendix that tests the proposed signal. For the actual implementation provided here, the signal we use comes from one of our own published anomalies—the monetary policy exposure (MPE) index from Ozdagli and Velikov (2020). The following report is the direct output of the tools that results from inputting the .csv file containing firm-date observations on MPE.

Online Appendix for Assaying Anomalies: Monetary Policy Exposure and the Cross Section of Stock Returns

Robert Novy-Marx Mihail Velikov

February 13, 2024

Abstract

This report studies the asset pricing implications of Monetary Policy Exposure (MPE), and its robustness in predicting returns in the cross-section of equities using the protocol proposed by Novy-Marx and Velikov (2023). A value-weighted long/short trading strategy based on MPE achieves an annualized gross (net) Sharpe ratio of 0.61 (0.48), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 43 (19) bps/month with a t-statistic of 4.74 (2.03), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (Size, Amihud's illiquidity, Past trading volume, Price, Efficient frontier index, Inst Own and Market to Book) is 40 bps/month with a t-statistic of 4.72.

1 Introduction

The following automatically generated report tests the asset pricing implications of Monetary Policy Exposure (MPE), and its robustness in predicting returns in the cross-section of equities. It is produced using the methodology of Novy-Marx and Velikov (2023), from input data consisting of firm-month observations for the proposed predictor.[1]

2 Signal diagnostics

Figure 1 plots descriptive statistics for the MPE signal. Panel A plots the time-series of the mean, median, and interquartile range for MPE. On average, the cross-sectional mean (median) MPE is 0.54 (0.52) over the 1975 to 2021 sample, where the starting date is determined by the availability of the input MPE data. The signal's interquartile range spans -6.63 to 1.60. Panel B of Figure 1 plots the time-series of the coverage of the MPE signal for the CRSP universe. On average, the MPE signal is available for 47.03% of CRSP names, which on average make up 69.97% of total market capitalization.

3 Does MPE predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on MPE using NYSE breaks. The first two lines of Panel A report monthly average excess returns for each of the five portfolios and for the long/short portfolio that buys the high MPE portfolio and sells the low MPE portfolio. The rest of Panel A reports the portfolios' monthly abnormal returns relative to the five most common factor models: the CAPM, the Fama and French (1993) three-factor model

[1] It used version v0.4.1 of the publicly available code repository at https://github.com/velikov-mihail/AssayingAnomalies. See more details at http://AssayingAnomalies.com.

(FF3) and its variation that adds momentum (FF4), the Fama and French (2015) five-factor model (FF5), and its variation that adds momentum factor used in Fama and French (2018) (FF6). The table shows that the long/short MPE strategy earns an average return of 0.64% per month with a t-statistic of 4.19. The annualized Sharpe ratio of the strategy is 0.61. The alphas range from 0.29% to 0.54% per month and have t-statistics exceeding 3.01 everywhere. The lowest alpha is with respect to the FF3 factor model.

Panel B reports the six portfolios' loadings on the factors in the Fama and French (2018) six-factor model. The long/short strategy's most significant loading is 0.95, with a t-statistic of 29.42 on the SMB factor. Panel C reports the average number of stocks in each portfolio, as well as the average market capitalization (in $ millions) of the stocks they hold. In an average month, the five portfolios have at least 292 stocks and an average market capitalization of at least $195 million.

Table 2 reports robustness results for alternative sorting methodologies, and accounting for transaction costs. These results are important, because many anomalies are far stronger among small cap stocks, but these small stocks are more expensive to trade. Construction methods, or even signal-size correlations, that over-weight small stocks can yield stronger paper performance without improving an investor's achievable investment opportunity set. Panel A reports gross returns and alphas for the long/short strategies made using various different protfolio constructions. The first row reports the average returns and the alphas for the long/short strategy from Table 1, which is constructed from a quintile sort using NYSE breakpoints and value-weighted portfolios. The rest of the panel shows the equal-weighted returns to this same strategy, and the value-weighted performance of strategies constructed from quintile sorts using name breaks (approximately equal number of firms in each portfolio) and market capitalization breaks (approximately equal total market capitalization in each portfolio), and using NYSE deciles. The average return is lowest for the quintile sort using cap breakpoints and value-weighted portfolios, and equals 41 bps/month with a t-statistics of 2.91. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-two exceed two, and for nineteen exceed three.

Panel B reports for these same strategies the average monthly net returns and the generalized net alphas of Novy-Marx and Velikov (2016). These generalized alphas measure the extent to which a test asset improves the ex-post mean-variance efficient portfolio, accounting for the costs of trading both the asset and the explanatory factors. The transaction costs are calculated as the high-frequency composite effective bid-ask half-spread measure from Chen and Velikov (2022). The net average returns reported in the first column range between 33-65bps/month. The lowest return, (33 bps/month), is achieved from the quintile sort using cap breakpoints and value-weighted portfolios, and has an associated t-statistic of 2.38. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the MPE trading strategy improves the achievable mean-variance efficient frontier spanned by the factor models in twenty-four cases, and significantly expands the achievable frontier in fifteen cases.

Table 3 provides direct tests for the role size plays in the MPE strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and MPE, as well as average returns and alphas for long/short trading MPE strategies within each size quintile. Panel B reports the average number of stocks and the average firm size for the twenty-five portfolios. Among the largest stocks (those with market capitalization greater than the 80^{th} NYSE percentile), the MPE strategy achieves an average return of 41 bps/month with a t-statistic of 2.90. Among these large cap stocks, the alphas for the MPE strategy relative to the five most common factor models range from 15 to 44 bps/month with t-statistics between 1.17 and 3.05.

4 How does MPE perform relative to the zoo?

Figure 2 puts the performance of MPE in context, showing the long/short strategy performance relative to other strategies in the "factor zoo." It shows Sharpe ratio histograms, both for gross and net returns (Panel A and B, respectively), for 207 documented anomalies in the zoo.[2] The vertical red line shows where the Sharpe ratio for the MPE strategy falls in the distribution. The MPE strategy's gross (net) Sharpe ratio of 0.61 (0.48) is greater than 96% (99%) of anomaly Sharpe ratios, respectively.

Figure 3 plots the growth of a \$1 invested in these same 207 anomaly trading strategies (gray lines), and compares those with the growth of a \$1 invested in the MPE strategy (red line).[3] Ignoring trading costs, a \$1 invested in the MPE strategy would have yielded \$23.09 which ranks the MPE strategy in the top 3% across the 207 anomalies. Accounting for trading costs, a \$1 invested in the MPE strategy would have yielded \$9.93 which ranks the MPE strategy in the top 1% across the 207 anomalies.

Figure 4 plots percentile ranks for the 207 anomaly trading strategies in terms of gross and Novy-Marx and Velikov (2016) net generalized alphas with respect to the CAPM, and the Fama-French three-, four-, five-, and six-factor models from Table 1, and indicates the ranking of the MPE relative to those. Panel A shows that the MPE strategy gross alphas fall between the 67 and 90 percentiles across the five factor models. Panel B shows that, accounting for trading costs, a large fraction of anomalies have not improved the investment opportunity set of an investor with access to the factor models over the 197501 to 202112 sample. For example, 49%

[2]The anomalies come from March, 2022 release of the Chen and Zimmermann (2022) open source asset pricing dataset.

[3]The figure assumes an initial investment of \$1 in T-bills and \$1 long/short in the two sides of the strategy. Returns are compounded each month, assuming, as in Detzel et al. (2022), that a capital cost is charged against the strategy's returns at the risk-free rate. This excess return corresponds more closely to the strategy's economic profitability.

(54%) of the 207 anomalies would not have improved the investment opportunity set for an investor having access to the Fama-French three-factor (six-factor) model. The MPE strategy has a positive net generalized alpha for five out of the five factor models. In these cases MPE ranks between the 69 and 91 percentiles in terms of how much it could have expanded the achievable investment frontier.

5 Does MPE add relative to related anomalies?

With so many anomalies, it is possible that any proposed, new cross-sectional predictor is just capturing some combination of known predictors. It is consequently natural to investigate to what extent the proposed predictor adds additional predictive power beyond the most closely related anomalies. Closely related anomalies are more likely to be formed on the basis of signals with higher absolute correlations. Figure 5 plots a name histogram of the correlations of MPE with 202 filtered anomaly signals.[4] Figure 6 also shows an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

A closely related anomaly is also more likely to price MPE or at least to weaken the power MPE has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of MPE conditioning on each of the 202 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{MPE} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{MPE} MPE_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 202 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{MPE,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 202 filtered anomaly trading strategies at a time. The strategies employed in the spanning tests are constructed using quintile sorts, value-

[4]When performing tests at the underlying signal level (e.g., the correlations plotted in Figure 5), we filter the 207 anomalies to avoid small sample issues. For each anomaly, we calculate the common stock observations in an average month for which both the anomaly and the test signal are available. In the filtered anomaly set, we drop anomalies with fewer than 100 common stock observations in an average month.

weighting, and NYSE breakpoints. Panel C plots t-statistics on the average returns to strategies constructed by conditional double sorts. In each month, we sort stocks into quintiles based one of the 202 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on MPE. Stocks are finally grouped into five MPE portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted MPE trading strategies conditioned on each of the 202 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on MPE and the six anomalies most closely-related to it. The six most-closely related anomalies are picked as those with the highest combined rank where the ranks are based on the absolute value of the Spearman correlations in Panel B of Figure 5 and the R^2 from the spanning tests in Figure 7, Panel B. Controlling for each of these signals at a time, the t-statistics on the MPE signal in these Fama-MacBeth regressions exceed 5.58, with the minimum t-statistic occurring when controlling for Inst Own and Market to Book. Controlling for all six closely related anomalies, the t-statistic on MPE is 4.69.

Similarly, Table 5 reports results from spanning tests that regress returns to the MPE strategy onto the returns of the six most closely-related anomalies and the six Fama-French factors. Controlling for the six most-closely related anomalies individually, the MPE strategy earns alphas that range from 37-43bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 4.25, which is achieved when controlling for Inst Own and Market to Book. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the MPE trading strategy achieves an alpha of 40bps/month with a t-statistic of 4.72.

6 Does MPE add relative to the whole zoo?

Finally, we can ask how much adding MPE to the entire factor zoo could improve investment performance. Figure 8 plots the growth of $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). The combinations use either the 147 anomalies from the zoo that satisfy our inclusion criteria (blue lines) or these 147 anomalies augmented with the MPE signal.[5] We consider six different methods for combining signals.

Panel A shows results using "Average rank" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are calculated on the basis of their average cross-sectional percentile rank across return predictors, and the predictors are all signed so that higher ranks are associated with higher average returns. For this method, $1 investment in the 147-anomaly combination strategy grows to $676.42, while $1 investment in the combination strategy that includes MPE grows to $545.45.

Panel B shows results using "Weighted-Average rank" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are calculated as weighted-average cross-sectional percentile rank across return predictors, and the predictors are all signed so that higher ranks are associated with higher average returns and the weights are determined by the average returns over the past ten years to the long/short strategies based on the individual signals. For this method, $1 investment in the 147-anomaly combination strategy grows to $39.74, while $1 investment in the combination strategy that includes MPE grows to $33.85.

Panel C shows results using "Fama-MacBeth" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are calculated from Fama and MacBeth (1973) regressions following Haugen and Baker (1996) and

[5] We filter the 207 Chen and Zimmermann (2022) anomalies and require for each anomaly the average month to have at least 40% of the cross-sectional observations available for market capitalization on CRSP in the period for which MPE is available.

Lewellen (2015) using only data in the investor's information set at the time of portfolio formation. The estimation uses rolling ten years of data, so the actual strategies begin ten years later for this combination method. For this method, \$1 investment in the 147-anomaly combination strategy grows to \$3961.83, while \$1 investment in the combination strategy that includes MPE grows to \$4184.71.

Panel D shows results using "Partial Least Squares" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are calculated from partial least squares (PLS) filtering procedure following Light et al. (2017) using only data in the investor's information set at the time of portfolio formation. The estimation uses rolling ten years of data, so the actual strategies begin ten years later for this combination method. For this method, \$1 investment in the 147-anomaly combination strategy grows to \$62.07, while \$1 investment in the combination strategy that includes MPE grows to \$82.85.

Panel E shows results using "IPCA" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are calculated from the instrumented principal component analysis (IPCA) procedure of Kelly et al. (2019) using only data in the investor's information set at the time of portfolio formation. The estimation uses rolling ten years of data, so the actual strategies begin ten years later for this combination method. For this method, \$1 investment in the 147-anomaly combination strategy grows to \$873.06, while \$1 investment in the combination strategy that includes MPE grows to \$550.06.

Panel F shows results using "LASSO" as the combination method. This method sorts stocks on the basis of forecast excess returns, where these are estimated by least absolute shrinkage and selection operator (LASSO) using only data in the investor's information set at the time of portfolio formation. Following Chen and Velikov (2022), LASSO penalty (λ) is selected by minimizing the mean squared error (MSE) estimated by 5-fold cross validation. The estimation uses rolling ten years of data, so the actual strategies begin ten years later for this combination method. For this method, \$1 investment in the 147-anomaly combination strategy grows to \$2019.46, while \$1 investment in the combination strategy that includes MPE grows to \$1489.58.

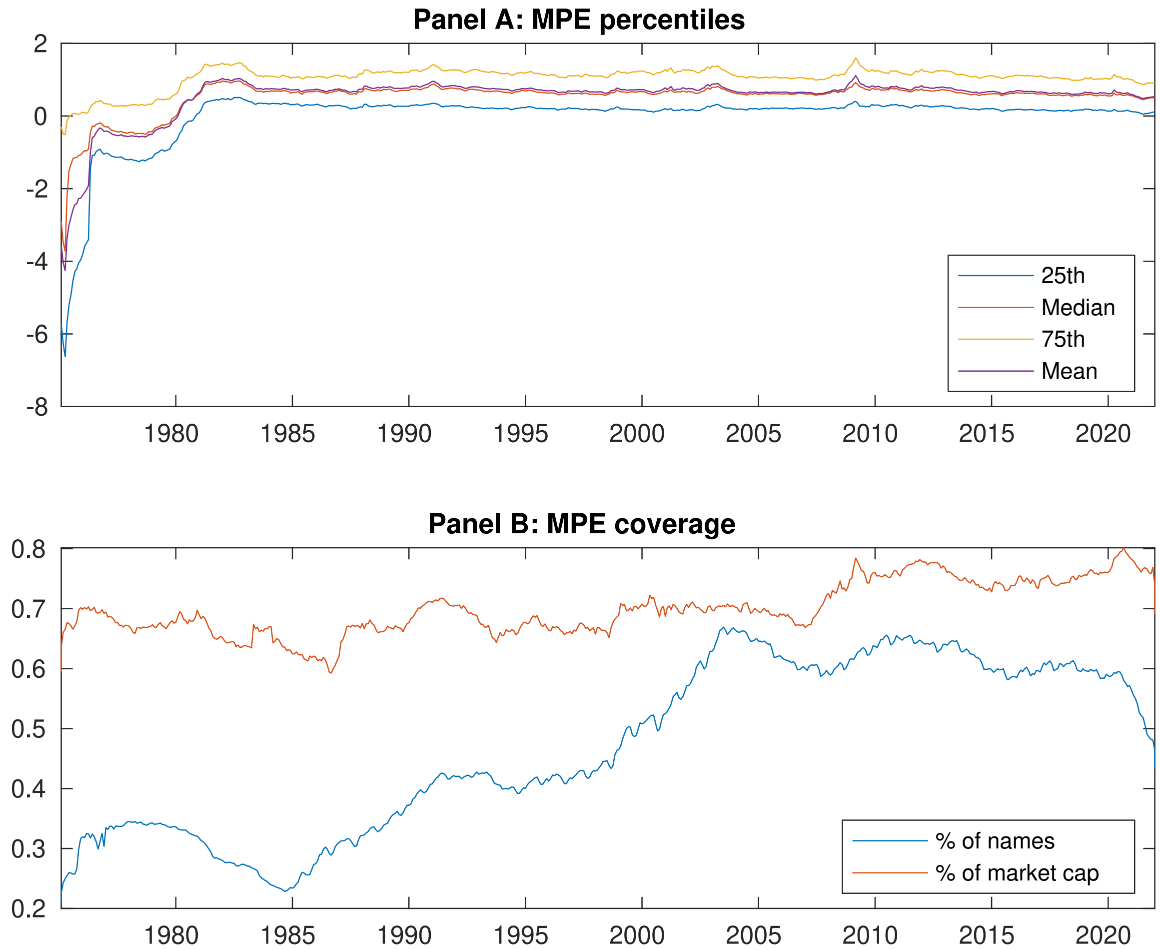


Figure 1: Times series of MPE percentiles and coverage.
This figure plots descriptive statistics for MPE. Panel A shows cross-sectional percentiles of MPE over the sample. Panel B plots the monthly coverage of MPE relative to the universe of CRSP stocks with available market capitalizations.

Table 1: Basic sort: VW, quintile, NYSE-breaks

This table reports average excess returns and alphas for portfolios sorted on MPE. At the end of each month, we sort stocks into five portfolios based on their signal using NYSE breakpoints. Panel A reports average value-weighted quintile portfolio (L,2,3,4,H) returns in excess of the risk-free rate, the long-short extreme quintile portfolio (H-L) return, and alphas with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor, Fama and French (2015) five-factor model, and the Fama and French (2015) five-factor model augmented with the Carhart (1997) momentum factor following Fama and French (2018). Panel B reports the factor loadings for the quintile portfolios and long-short extreme quintile portfolio in the Fama and French (2015) five-factor model. Panel C reports the average number of stocks and market capitalization of each portfolio. T-statistics are in brackets. The sample period is 197501 to 202112.

Panel A: Excess returns and alphas on MPE-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.66	0.74	0.98	1.04	1.30	0.64
	[3.45]	[3.79]	[4.50]	[4.45]	[5.35]	[4.19]
α_{CAPM}	-0.05	0.02	0.19	0.22	0.49	0.54
	[-0.97]	[0.40]	[2.58]	[2.27]	[3.98]	[3.54]
α_{FF3}	0.03	-0.00	0.14	0.11	0.32	0.29
	[0.81]	[-0.01]	[1.97]	[1.39]	[3.97]	[3.01]
α_{FF4}	0.03	0.03	0.21	0.22	0.48	0.45
	[0.67]	[0.42]	[2.98]	[2.85]	[6.44]	[4.99]
α_{FF5}	-0.01	-0.06	0.09	0.09	0.29	0.30
	[-0.35]	[-1.06]	[1.26]	[1.15]	[3.49]	[3.09]
α_{FF6}	-0.01	-0.04	0.15	0.18	0.41	0.43
	[-0.34]	[-0.63]	[2.12]	[2.34]	[5.67]	[4.74]
Panel B: Fama and French (2018) 6-factor model loadings for MPE-sorted portfolios						
β_{MKT}	1.00	1.00	1.07	1.06	1.02	0.02
	[98.48]	[69.38]	[63.26]	[59.02]	[58.57]	[0.93]
β_{SMB}	-0.14	0.08	0.24	0.54	0.81	0.95
	[-9.50]	[3.59]	[9.49]	[19.85]	[30.86]	[29.42]
β_{HML}	-0.22	-0.02	-0.02	0.08	0.05	0.27
	[-11.66]	[-0.89]	[-0.69]	[2.34]	[1.42]	[6.65]
β_{RMW}	0.06	0.11	0.10	0.13	0.16	0.10
	[2.88]	[3.78]	[3.03]	[3.79]	[4.66]	[2.41]
β_{CMA}	0.09	0.12	0.13	0.00	0.14	0.05
	[3.06]	[2.94]	[2.66]	[0.09]	[2.73]	[0.77]
β_{UMD}	-0.00	-0.04	-0.10	-0.14	-0.21	-0.21
	[-0.08]	[-3.10]	[-6.05]	[-8.16]	[-12.67]	[-10.20]
Panel C: Average number of firms (n) and market capitalization (me)						
n	304	292	341	456	929	
me ($\$10^6$)	4823	1790	759	376	195	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the MPE strategy construction methodology. In each panel, the first row shows results from a quintile, value-weighted sort using NYSE break points as employed in Table 1. Each of the subsequent rows deviates in one of the three choices at a time, and the choices are specified in the first three columns. For each strategy construction methodology, the table reports average excess returns and alphas with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor, Fama and French (2015) five-factor model, and the Fama and French (2015) five-factor model augmented with the Carhart (1997) momentum factor following Fama and French (2018). Panel A reports average returns and alphas with no adjustment for trading costs. Panel B reports net average returns and Novy-Marx and Velikov (2016) generalized alphas as prescribed by Detzel et al. (2022). T-statistics are in brackets. The sample period is 197501 to 202112.

Panel A: Gross Returns and Alphas								
Portfolios	Breaks	Weights	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Quintile	NYSE	VW	0.64	0.54	0.29	0.45	0.30	0.43
			[4.19]	[3.54]	[3.01]	[4.99]	[3.09]	[4.74]
Quintile	NYSE	EW	1.06	1.13	0.97	1.08	0.89	0.98
			[8.71]	[9.28]	[9.33]	[10.39]	[8.42]	[9.45]
Quintile	Name	VW	0.89	0.77	0.50	0.75	0.51	0.71
			[5.04]	[4.40]	[4.07]	[6.73]	[4.13]	[6.39]
Quintile	Cap	VW	0.41	0.40	0.16	0.24	0.03	0.10
			[2.91]	[2.80]	[1.40]	[2.08]	[0.25]	[0.91]
Decile	NYSE	VW	0.79	0.68	0.37	0.57	0.31	0.48
			[4.37]	[3.76]	[3.13]	[5.18]	[2.63]	[4.39]
Panel B: Net Returns and Novy-Marx and Velikov (2016) generalized alphas								
Portfolios	Breaks	Weights	r^e_{net}	α^*_{CAPM}	α^*_{FF3}	α^*_{FF4}	α^*_{FF5}	α^*_{FF6}
Quintile	NYSE	VW	0.50	0.39	0.17	0.25	0.10	0.19
			[3.27]	[2.53]	[1.76]	[2.79]	[1.02]	[2.03]
Quintile	NYSE	EW	0.63	0.69	0.55	0.60	0.43	0.49
			[5.16]	[5.65]	[5.29]	[5.89]	[4.13]	[4.76]
Quintile	Name	VW	0.65	0.52	0.28	0.41	0.22	0.35
			[3.67]	[2.97]	[2.31]	[3.74]	[1.78]	[3.05]
Quintile	Cap	VW	0.33	0.32	0.12	0.16		0.01
			[2.38]	[2.23]	[1.08]	[1.47]		[0.06]
Decile	NYSE	VW	0.57	0.46	0.19	0.29	0.06	0.17
			[3.17]	[2.50]	[1.62]	[2.63]	[0.46]	[1.50]

Table 3: Conditional sort on size and MPE

This table presents results for conditional double sorts on size and MPE. In each month, stocks are first sorted into quintiles based on size using NYSE breakpoints. Then, within each size quintile, stocks are further sorted based on MPE. Finally, they are grouped into twenty-five portfolios based on the intersection of the two sorts. Panel A presents the average returns to the 25 portfolios, as well as strategies that go long stocks with high MPE and short stocks with low MPE .Panel B documents the average number of firms and the average firm size for each portfolio. The sample period is 197501 to 202112.

Panel A: portfolio average returns and time-series regression results

		MPE Quintiles					MPE Strategies					
		(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Size quintiles	(1)	0.52	0.97	1.23	1.51	1.95	1.43	1.61	1.52	1.53	1.25	1.28
		[1.57]	[3.44]	[4.56]	[5.46]	[6.92]	[9.26]	[10.72]	[10.63]	[10.52]	[8.92]	[9.08]
	(2)	0.56	0.93	1.07	1.18	1.45	0.89	1.01	0.96	0.92	0.72	0.71
		[1.79]	[3.44]	[4.08]	[4.49]	[5.44]	[6.54]	[7.52]	[7.47]	[7.06]	[5.97]	[5.84]
	(3)	0.60	0.86	1.07	1.01	1.24	0.64	0.73	0.66	0.65	0.42	0.44
		[2.19]	[3.46]	[4.37]	[4.12]	[5.03]	[4.96]	[5.70]	[5.22]	[5.04]	[3.47]	[3.54]
	(4)	0.61	0.81	0.97	0.98	1.09	0.48	0.55	0.44	0.47	0.34	0.37
		[2.38]	[3.59]	[4.26]	[4.21]	[4.59]	[3.95]	[4.54]	[3.82]	[4.03]	[2.91]	[3.17]
	(5)	0.64	0.54	0.72	0.80	1.05	0.41	0.44	0.24	0.29	0.15	0.20
		[3.04]	[2.89]	[3.74]	[4.23]	[5.12]	[2.90]	[3.05]	[1.87]	[2.27]	[1.17]	[1.55]

Panel B: Portfolio average number of firms and market capitalization

		MPE Quintiles					MPE Quintiles				
		Average n					Average market capitalization ($\$10^6$)				
		(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
Size quintiles	(1)	236	237	236	236	236	29	30	26	19	11
	(2)	77	77	77	77	77	44	45	43	41	39
	(3)	56	56	56	56	56	77	76	73	71	68
	(4)	49	49	49	49	49	175	171	160	155	144
	(5)	45	45	45	45	45	1967	1598	1189	953	740

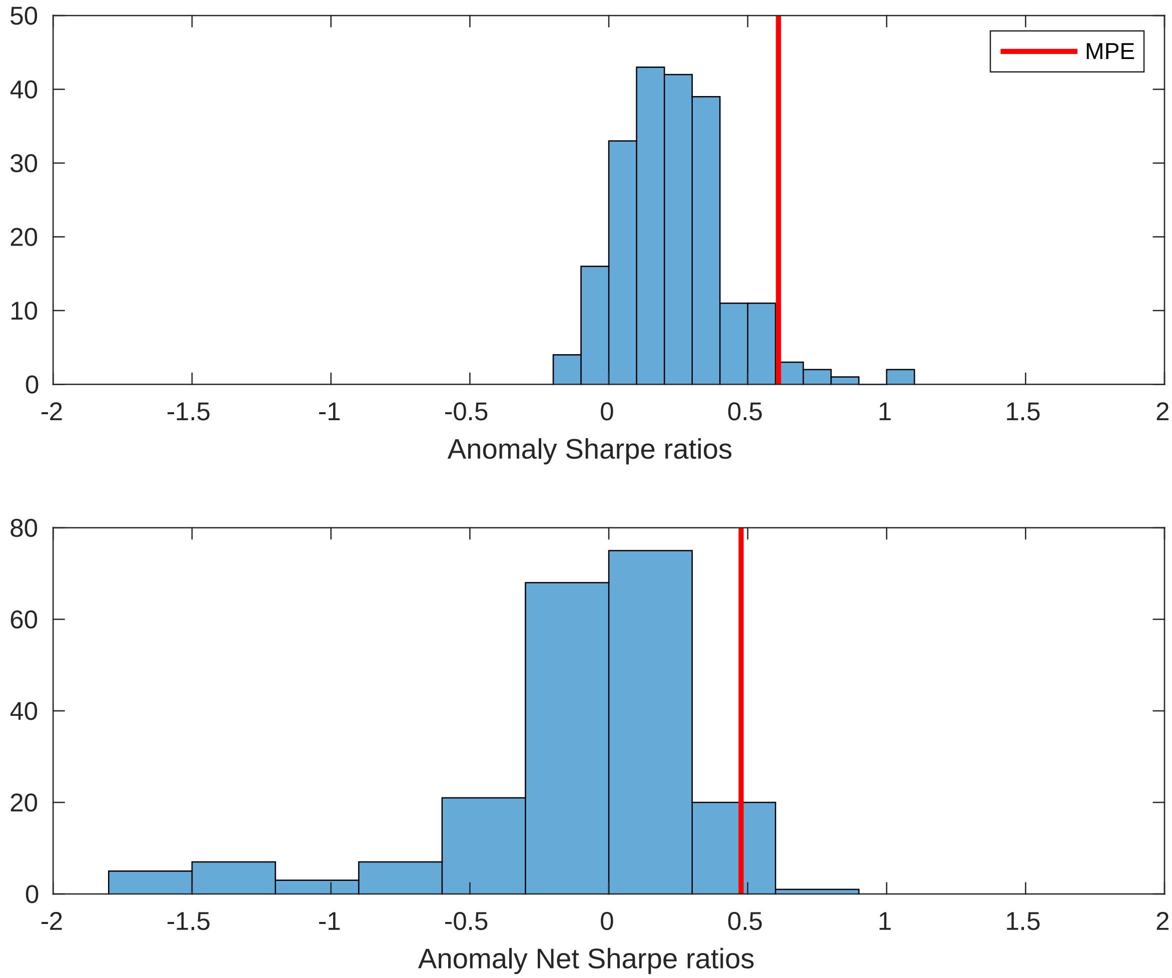


Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 207 anomalies, and compares the Sharpe ratio of the MPE with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

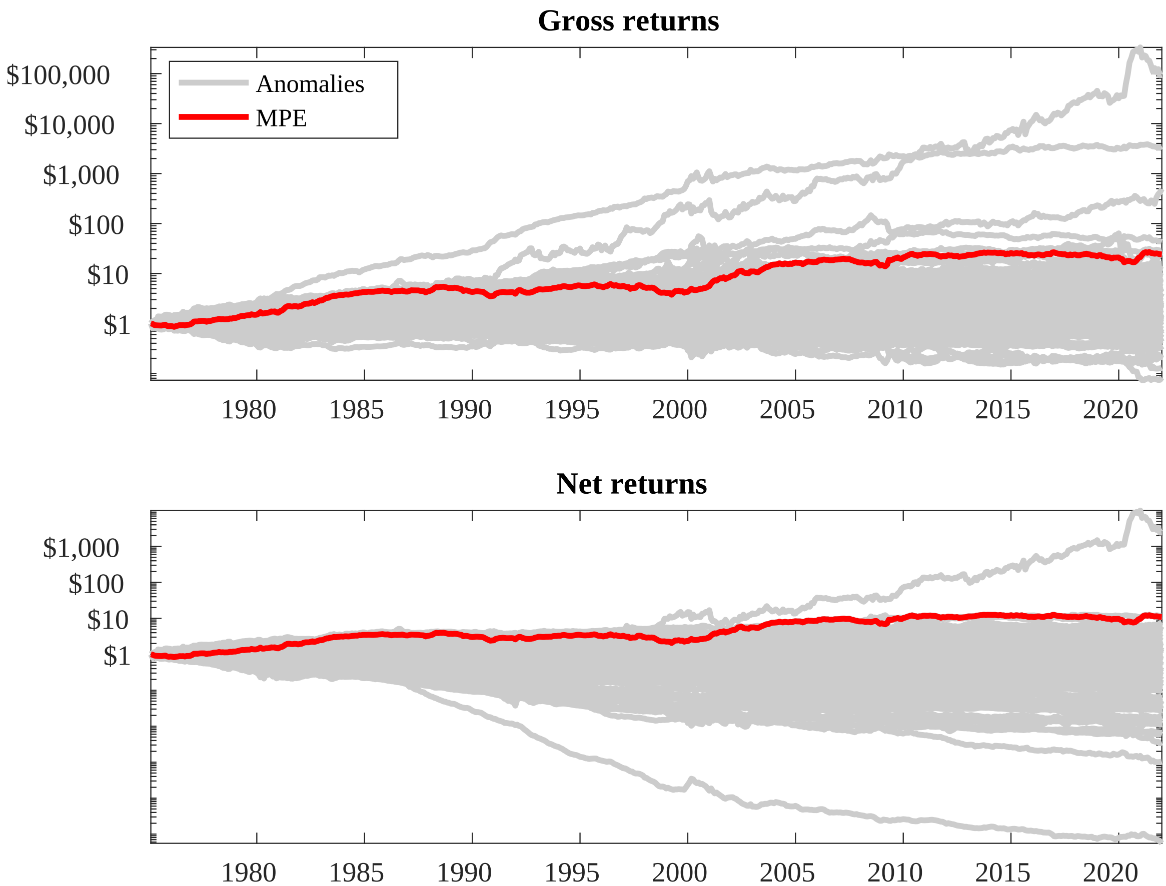
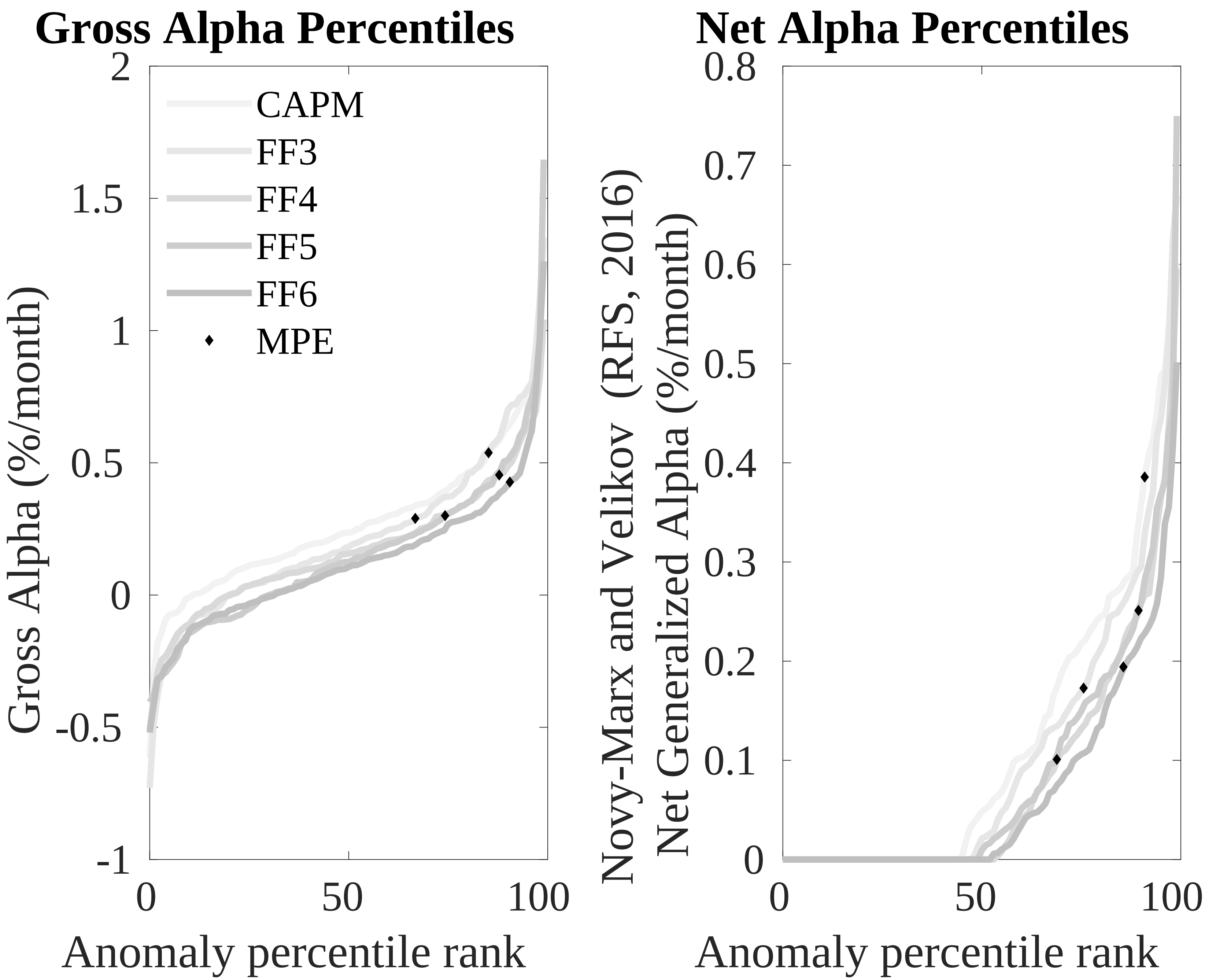


Figure 3: Dollar invested.
This figure plots the growth of a $1 invested in 207 anomaly trading strategies (gray lines), and compares those with the MPE trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross strategy returns. Panel B plots results for net strtaegy returns.


Gross Alpha Percentiles
Net Alpha Percentiles
CAPM
FF3
FF4
FF5
FF6
MPE
Gross Alpha (%/month)
Novy-Marx and Velikov (RFS, 2016)
Net Generalized Alpha (%/month)
Anomaly percentile rank
Anomaly percentile rank
2
1.5
1
0.5
0
-0.5
-1
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
0
50
100

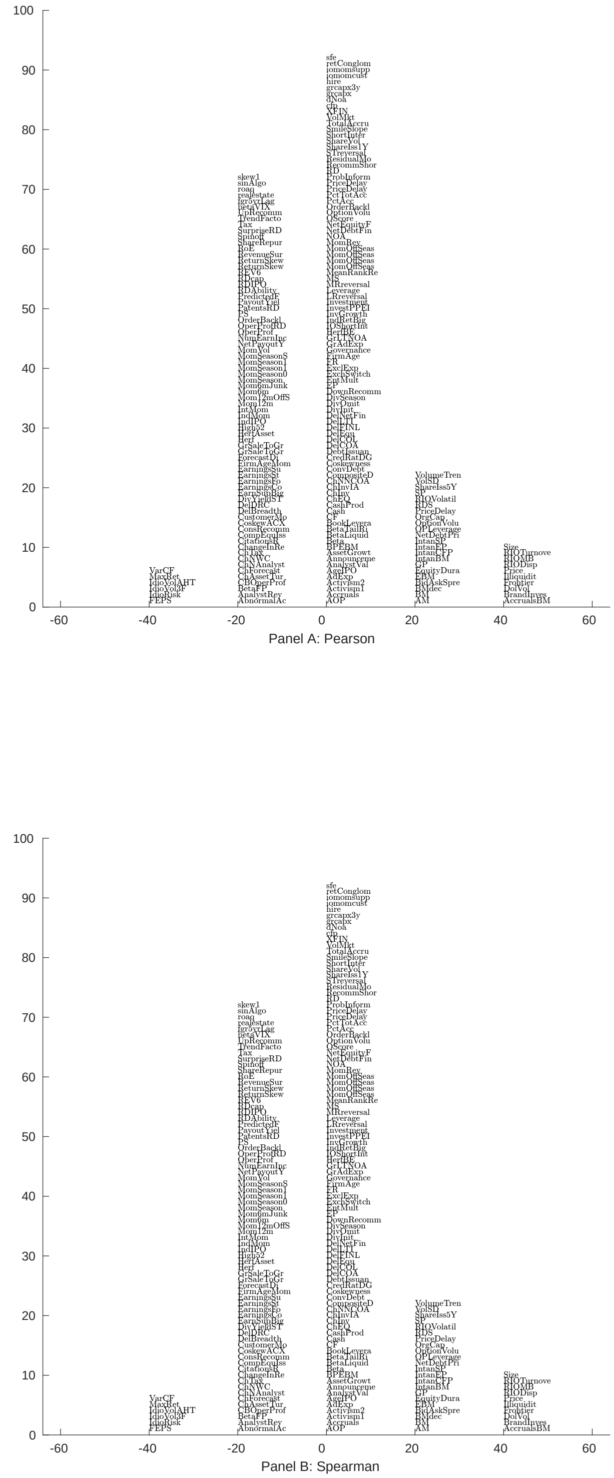


Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 202 filtered anomaly signals with MPE. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

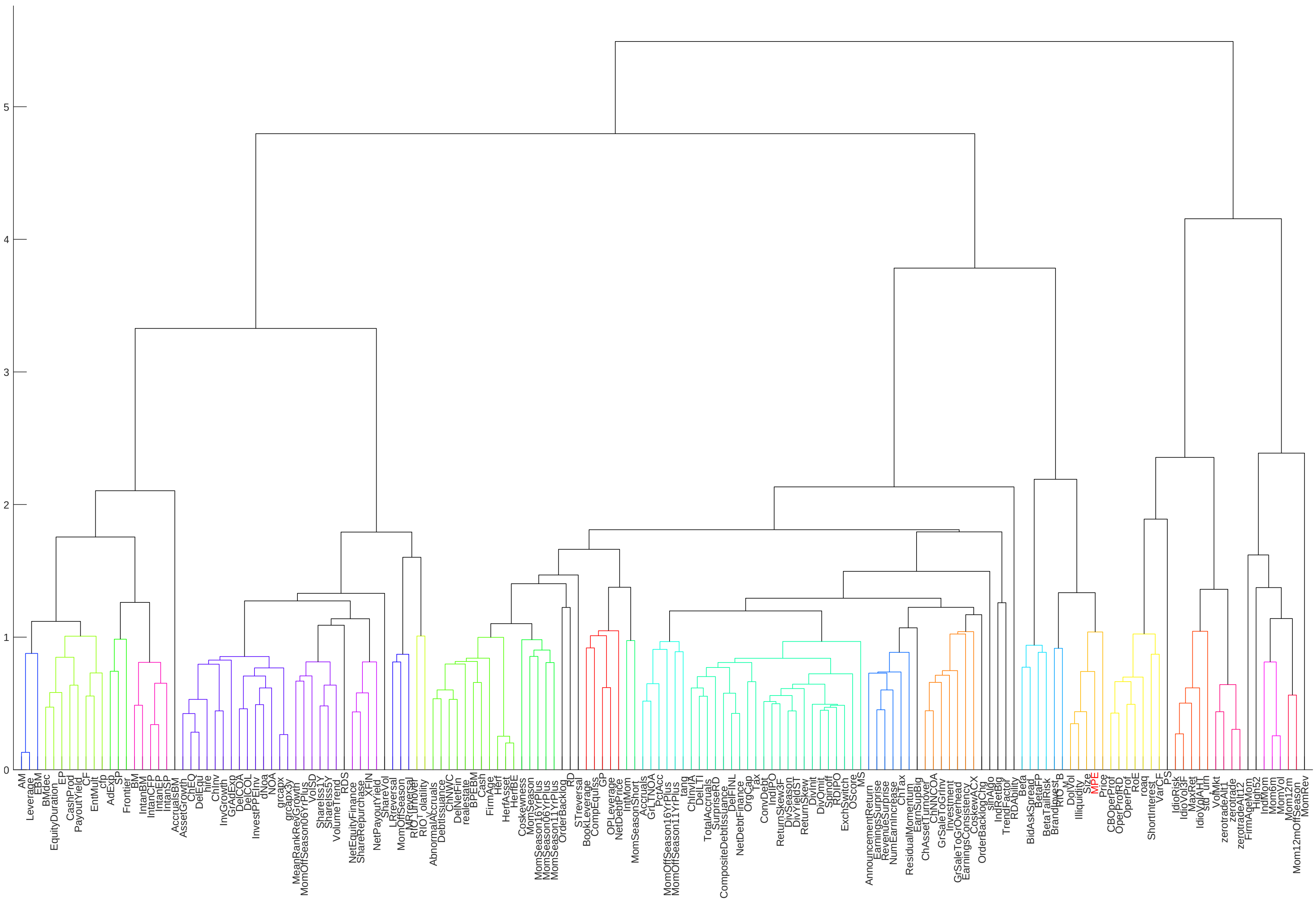


Figure 6: Agglomerative hierarchical cluster plot
This figure plots an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

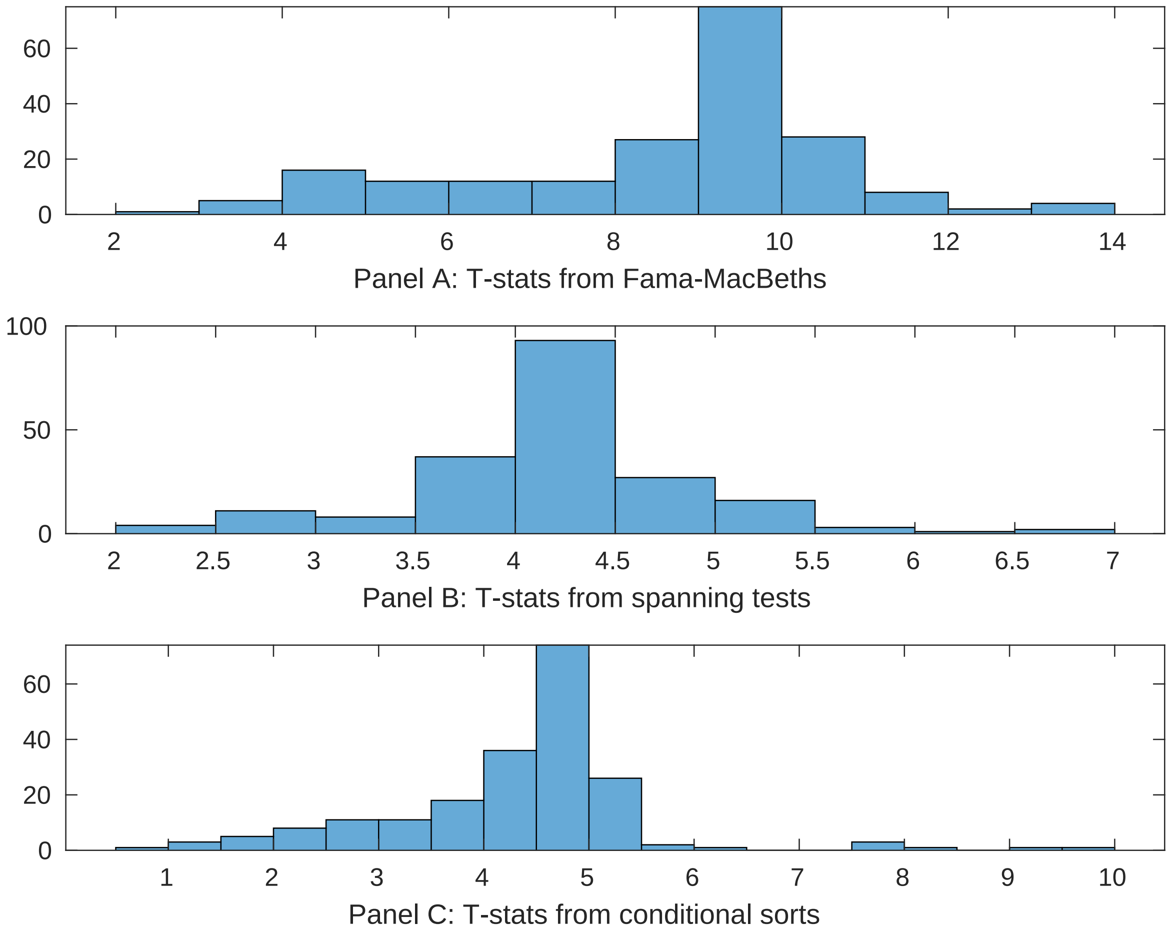


Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of MPE conditioning on each of the 202 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{MPE} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{MPE}MPE_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 202 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{MPE,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 202 filtered anomaly trading strategies at a time. The strategies employed in the spanning tests are constructed using quintile sorts, value-weighting, and NYSE breakpoints. Panel C plots t-statistics on the average returns to strategies constructed by conditional double sorts. In each month, we sort stocks into quintiles based one of the 202 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on MPE. Stocks are finally grouped into five MPE portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted MPE trading strategies conditioned on each of the 202 filtered anomalies.

Table 4: Fama-MacBeths controlling for most closely related anomalies

This table presents Fama-MacBeth results of returns on MPE. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{MPE} MPE_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Size, Amihud's illiquidity, Past trading volume, Price, Efficient frontier index, Inst Own and Market to Book. These anomalies were picked as those with the lowest absolute sum of t-statistics across the three Panels in Figure 7. The sample period is 197501 to 202112.

Intercept	0.89	0.98	0.73	0.87	0.11	0.10	0.13
	[3.26]	[3.89]	[2.59]	[2.65]	[4.16]	[4.67]	[0.29]
MPE	0.80	0.63	0.85	0.80	0.60	0.62	0.76
	[11.49]	[9.61]	[12.61]	[12.34]	[8.41]	[5.58]	[4.69]
Anomaly 1	0.37						0.90
	[0.93]						[2.28]
Anomaly 2		0.19					-0.65
		[2.68]					[-1.13]
Anomaly 3			-0.31				-0.72
			[-1.02]				[-1.01]
Anomaly 4				0.16			-0.74
				[0.00]			[-1.66]
Anomaly 5					0.24		-0.83
					[2.81]		[-0.84]
Anomaly 6						0.18	0.92
						[0.31]	[1.35]
# months	563	563	563	563	563	558	477
$\bar{R}^2$(%)	1	1	1	2	1	2	0

Table 5: Spanning tests controlling for most closely related anomalies
This table presents spanning tests results of regressing returns to the MPE trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{MPE} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, where X_k indicates each of the six most-closely related anomalies and f_j indicates the six factors from the Fama and French (2015) five-factor model augmented with the Carhart (1997) momentum factor. The six most closely related anomalies, X, are Size, Amihud's illiquidity, Past trading volume, Price, Efficient frontier index, Inst Own and Market to Book. These anomalies were picked as those with the highest combined rank where the ranks are based on the absolute value of the Spearman correlations in Panel B of Figure 5 and the R^2 from the spanning tests in Figure 7, Panel B. The sample period is 197501 to 202112.

Intercept	0.42	0.43	0.43	0.42	0.37	0.43	0.40
	[4.67]	[4.75]	[4.77]	[4.64]	[4.25]	[4.86]	[4.72]
Anomaly 1	17.87						50.61
	[3.33]						[5.41]
Anomaly 2		-13.08					-32.18
		[-1.75]					[-2.86]
Anomaly 3			-9.04				-21.32
			[-1.42]				[-2.09]
Anomaly 4				8.68			-11.46
				[2.42]			[-2.41]
Anomaly 5					17.95		14.66
					[7.12]		[5.54]
Anomaly 6						13.76	13.50
						[3.99]	[4.06]
mkt	2.48	0.95	0.51	0.39	1.02	0.98	-2.62
	[1.16]	[0.42]	[0.21]	[0.17]	[0.49]	[0.46]	[-1.07]
smb	73.31	109.74	103.63	85.74	86.15	79.85	81.85
	[10.17]	[12.00]	[14.77]	[17.35]	[25.85]	[16.21]	[8.71]
hml	25.43	28.98	29.17	26.19	19.86	28.55	29.84
	[6.29]	[6.88]	[6.71]	[6.47]	[4.96]	[7.10]	[6.99]
rmw	12.43	7.85	10.56	14.52	9.87	15.81	12.05
	[2.96]	[1.80]	[2.52]	[3.19]	[2.47]	[3.62]	[2.44]
cma	4.95	6.30	5.70	3.88	2.74	6.66	12.62
	[0.80]	[1.00]	[0.91]	[0.62]	[0.46]	[1.07]	[2.11]
umd	-17.87	-21.58	-22.52	-16.22	-10.26	-23.73	-14.71
	[-7.68]	[-10.30]	[-10.04]	[-5.44]	[-4.03]	[-11.04]	[-4.72]
# months	563	563	563	563	563	563	563
$\bar{R}^2(\%)$	70	69	69	70	72	70	74

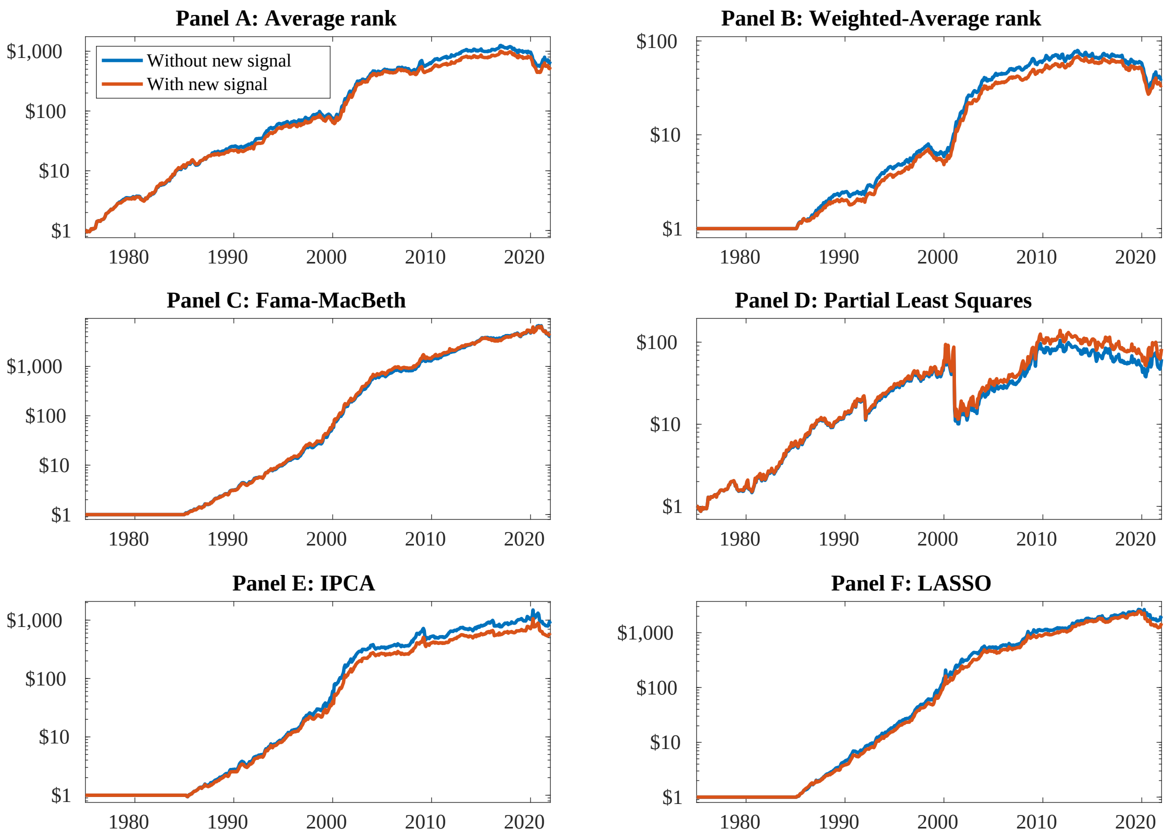


Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 147 anomalies. The red solid lines indicate combination trading strategies that utilize the 147 anomalies as well as MPE. Panel A shows results using "Average rank" as the combination method. Panel B shows results using "Weighted-Average rank" as the combination method. Panel C shows results using "Fama-MacBeth" as the combination method. Panel D shows results using "Partial Least Squares" as the combination method. Panel E shows results using "IPCA" as the combination method. Panel F shows results using "LASSO" as the combination method. See Section 6 for details on the combination methods.

References

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

Chen, A. and Velikov, M. (2022). Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis*, Forthcoming.

Chen, A. Y. and Zimmermann, T. (2022). Open source cross-sectional asset pricing. *Critical Finance Review*, 27(2):207–264.

Detzel, A., Novy-Marx, R., and Velikov, M. (2022). Model comparison with transaction costs. *Journal of Finance, Forthcoming*.

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

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

Fama, E. F. and French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, 128(2):234–252.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy*, 81(3):607–636.

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

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

Lewellen, J. (2015). The cross-section of expected returns. *Critical Finance Review*, 4(1):1–44.

Light, N., Maslov, D., and Rytchkov, O. (2017). Aggregation of information about the cross section of stock returns: A latent variable approach. *Review of Financial Studies*, 30:1339–1381.

Novy-Marx, R. and Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies*, 29(1):104–147.

Novy-Marx, R. and Velikov, M. (2023). Assaying anomalies. *Working paper*.

Ozdagli, A. and Velikov, M. (2020). Show me the money: The monetary policy risk premium. *Journal of Financial Economics*, 135(2):320–339.