Equity Dilution Factor and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Equity Dilution Factor (EDF), 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 EDF achieves an annualized gross (net) Sharpe ratio of 0.61 (0.54), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 20 (20) bps/month with a t-statistic of 2.68 (2.72), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (Share issuance (1 year), Growth in book equity, Net Payout Yield, Share issuance (5 year), Change in equity to assets, Asset growth) is 16 bps/month with a t-statistic of 2.39.

1 Introduction

Market efficiency remains a central question in asset pricing, with researchers continually seeking to understand how information is incorporated into security prices. While extensive literature documents various return predictors in the cross-section of stocks, debate persists about which signals genuinely capture mispricing or risk premiums versus those that are merely statistical artifacts. A particularly important yet understudied channel of information flow concerns how firms' equity issuance and dilution decisions affect future returns.

Prior research has examined various aspects of equity issuance, including seasoned equity offerings, share repurchases, and general changes in shares outstanding. However, these studies typically focus on discrete corporate events or aggregate measures that may not fully capture the nuanced ways in which firms' ongoing equity decisions affect shareholder value. This gap is notable given that changes in equity structure represent one of management's most direct ways to impact existing shareholders' claims on firm assets.

We propose that systematic tracking of equity dilution provides valuable information about future stock returns through several economic channels. First, following Myers and Majluf (1984), managers who act in existing shareholders' interests should issue equity primarily when shares are overvalued, suggesting that equity dilution may signal overvaluation. Second, building on Baker and Wurgler (2002), the market timing hypothesis predicts that firms opportunistically issue equity when the cost of equity is unusually low, implying subsequent underperformance as prices revert to fundamental values.

The information content of equity dilution may be particularly valuable because it represents an aggregation of multiple corporate decisions that affect shareholders' claims. While Pontiff and Woodgate (2008) show that changes in shares outstanding predict returns, our Equity Dilution Factor (EDF) provides a more comprehensive measure by incorporating both direct share issuance and indirect forms of dilution such as option exercises and convertible securities. This broader perspective aligns with Eckbo and Masulis (1995)'s framework showing how various forms of equity issuance contain different degrees of information.

Moreover, building on Daniel and Titman (1999), we hypothesize that the market may underreact to the cumulative impact of incremental equity dilution, as investors face cognitive constraints in aggregating multiple small changes in equity claims. This suggests that a systematic factor capturing total dilution effects could identify stocks likely to underperform as the market gradually incorporates this information.

Our empirical analysis reveals that the Equity Dilution Factor (EDF) strongly predicts cross-sectional stock returns. A value-weighted long-short strategy based on EDF quintiles generates a monthly alpha of 20 basis points (t-statistic = 2.68) relative to the Fama-French six-factor model. The strategy's economic significance is substantial, achieving an annualized Sharpe ratio of 0.61 before trading costs and 0.54 after accounting for transaction costs.

Importantly, EDF's predictive power remains robust across various methodological specifications and controls. The signal maintains significance even among large-cap stocks, with the long-short strategy earning a monthly alpha of 24-31 basis points (t-statistics between 2.55 and 3.44) in the largest size quintile. This finding addresses common concerns about anomalies being driven primarily by small, illiquid stocks.

Further tests demonstrate that EDF's predictive ability is distinct from known factors. Controlling for the six most closely related anomalies and the Fama-French six factors simultaneously, the EDF strategy still generates a monthly alpha of 16 basis points (t-statistic = 2.39). This persistence suggests that EDF captures unique information not reflected in existing factors.

Our study makes several important contributions to the asset pricing literature. First, we extend the equity issuance literature pioneered by Loughran and Ritter

(1995) and Pontiff and Woodgate (2008) by developing a more comprehensive measure that captures both direct and indirect forms of equity dilution. Unlike previous work focused on specific corporate events, our EDF measure provides a unified framework for understanding how various equity-related decisions collectively impact future returns.

Second, we contribute to the growing literature on factor investing by introducing a novel predictor that is both economically intuitive and empirically robust. Our findings complement recent work by Hou et al. (2020) on factor replication by showing that EDF represents a distinct source of predictability that cannot be explained by existing factors. The signal's strong performance among large-cap stocks and after controlling for transaction costs addresses key criticisms of many previously documented anomalies.

Finally, our results have important implications for both academic research and investment practice. For academics, we provide new evidence on how markets process information about changes in equity structure, supporting theories of gradual information incorporation. For practitioners, our findings suggest profitable trading strategies that remain viable after accounting for implementation costs, with particular relevance for institutional investors focused on large-cap stocks.

2 Data

Our study investigates the predictive power of a financial signal derived from accounting data for cross-sectional returns, focusing specifically on equity dilution. We obtain accounting and financial data from COMPUSTAT, covering firm-level observations for publicly traded companies. To construct our signal, we use COMPUSTAT's item CSTK for common stock and item AO for shares outstanding. Common stock (CSTK) represents the total value of common shares issued by the company, while shares outstanding (AO) reflects the total number of shares currently held by shareholders. The construction of our Equity Dilution Factor follows a specific methodology where we calculate the year-over-year change in CSTK and scale it by the previous year's shares outstanding (AO). This calculation can be expressed as (CSTKt - CSTKt-1) / AOt-1, where t denotes the current period and t-1 represents the previous period. This signal captures the relative magnitude of changes in common stock issuance normalized by the existing share base, providing insight into the degree of equity dilution experienced by current shareholders. By focusing on this relationship, the signal aims to reflect aspects of capital structure changes and financing decisions in a manner that is both scalable and interpretable. We construct this measure using end-of-fiscal-year values for both CSTK and AO to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the EDF signal. Panel A plots the time-series of the mean, median, and interquartile range for EDF. On average, the cross-sectional mean (median) EDF is -1.23 (-0.00) over the 1966 to 2023 sample, where the starting date is determined by the availability of the input EDF data. The signal's interquartile range spans -0.42 to 0.00. Panel B of Figure 1 plots the time-series of the coverage of the EDF signal for the CRSP universe. On average, the EDF signal is available for 6.30% of CRSP names, which on average make up 7.80% of total market capitalization.

4 Does EDF predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on EDF 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 EDF portfolio and sells the low EDF 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 (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 EDF strategy earns an average return of 0.34% per month with a t-statistic of 4.61. The annualized Sharpe ratio of the strategy is 0.61. The alphas range from 0.20% to 0.36% per month and have t-statistics exceeding 2.68 everywhere. The lowest alpha is with respect to the FF6 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 0.29, with a t-statistic of 5.90 on the CMA 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 551 stocks and an average market capitalization of at least $1,405 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 decile sort using NYSE breakpoints and value-weighted portfolios, and equals 29 bps/month with a t-statistics of 3.27. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-four exceed two, and for seventeen 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 25-31bps/month. The lowest return, (25 bps/month), is achieved from the decile sort using NYSE breakpoints and value-weighted portfolios, and has an associated t-statistic of 2.82. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the EDF trading strategy improves the achievable mean-variance efficient frontier spanned by the factor models in twenty-five cases, and significantly expands the achievable frontier in twenty-two cases.

Table 3 provides direct tests for the role size plays in the EDF strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and EDF, as well as average returns and alphas for long/short trading EDF 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 EDF strategy achieves an average return of 31 bps/month with a t-statistic of 3.48. Among these large cap stocks, the alphas for the EDF strategy relative to the five most common factor models range from 24 to 31 bps/month with t-statistics between 2.55 and 3.44.

5 How does EDF perform relative to the zoo?

Figure 2 puts the performance of EDF 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 212 documented anomalies in the zoo.[1] The vertical red line shows where the Sharpe ratio for the EDF strategy falls in the distribution. The EDF strategy's gross (net) Sharpe ratio of 0.61 (0.54) is greater than 96% (99%) of anomaly Sharpe ratios, respectively.

Figure 3 plots the growth of a \$1 invested in these same 212 anomaly trading strategies (gray lines), and compares those with the growth of a \$1 invested in the EDF strategy (red line).[2] Ignoring trading costs, a \$1 invested in the EDF strategy would have yielded \$8.37 which ranks the EDF strategy in the top 1% across the 212 anomalies. Accounting for trading costs, a \$1 invested in the EDF strategy would have yielded \$6.30 which ranks the EDF strategy in the top 0% across the 212 anomalies.

Figure 4 plots percentile ranks for the 212 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 EDF relative to those. Panel A shows that

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

[2] 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.

the EDF strategy gross alphas fall between the 64 and 72 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 196606 to 202306 sample. For example, 45% (53%) of the 212 anomalies would not have improved the investment opportunity set for an investor having access to the Fama-French three-factor (six-factor) model. The EDF strategy has a positive net generalized alpha for five out of the five factor models. In these cases EDF ranks between the 84 and 89 percentiles in terms of how much it could have expanded the achievable investment frontier.

6 Does EDF 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 EDF with 210 filtered anomaly signals.[3] 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 EDF or at least to weaken the power EDF has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of EDF conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{EDF} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{EDF} EDF_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X

[3] When performing tests at the underlying signal level (e.g., the correlations plotted in Figure 5), we filter the 212 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.

stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{EDF,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 210 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 210 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on EDF. Stocks are finally grouped into five EDF portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted EDF trading strategies conditioned on each of the 210 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on EDF 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 EDF signal in these Fama-MacBeth regressions exceed 1.11, with the minimum t-statistic occurring when controlling for Net Payout Yield. Controlling for all six closely related anomalies, the t-statistic on EDF is 1.16.

Similarly, Table 5 reports results from spanning tests that regress returns to the EDF 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 EDF strategy earns alphas that range from 17-22bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.33, which is achieved when controlling for Net Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the EDF trading strategy achieves an alpha of 16bps/month with a t-statistic of 2.39.

7 Does EDF add relative to the whole zoo?

Finally, we can ask how much adding EDF 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 155 anomalies from the zoo that satisfy our inclusion criteria (blue lines) or these 155 anomalies augmented with the EDF signal.[4] We consider one 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 155-anomaly combination strategy grows to $2333.55, while $1 investment in the combination strategy that includes EDF grows to $2413.89.

[4] 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 EDF is available.

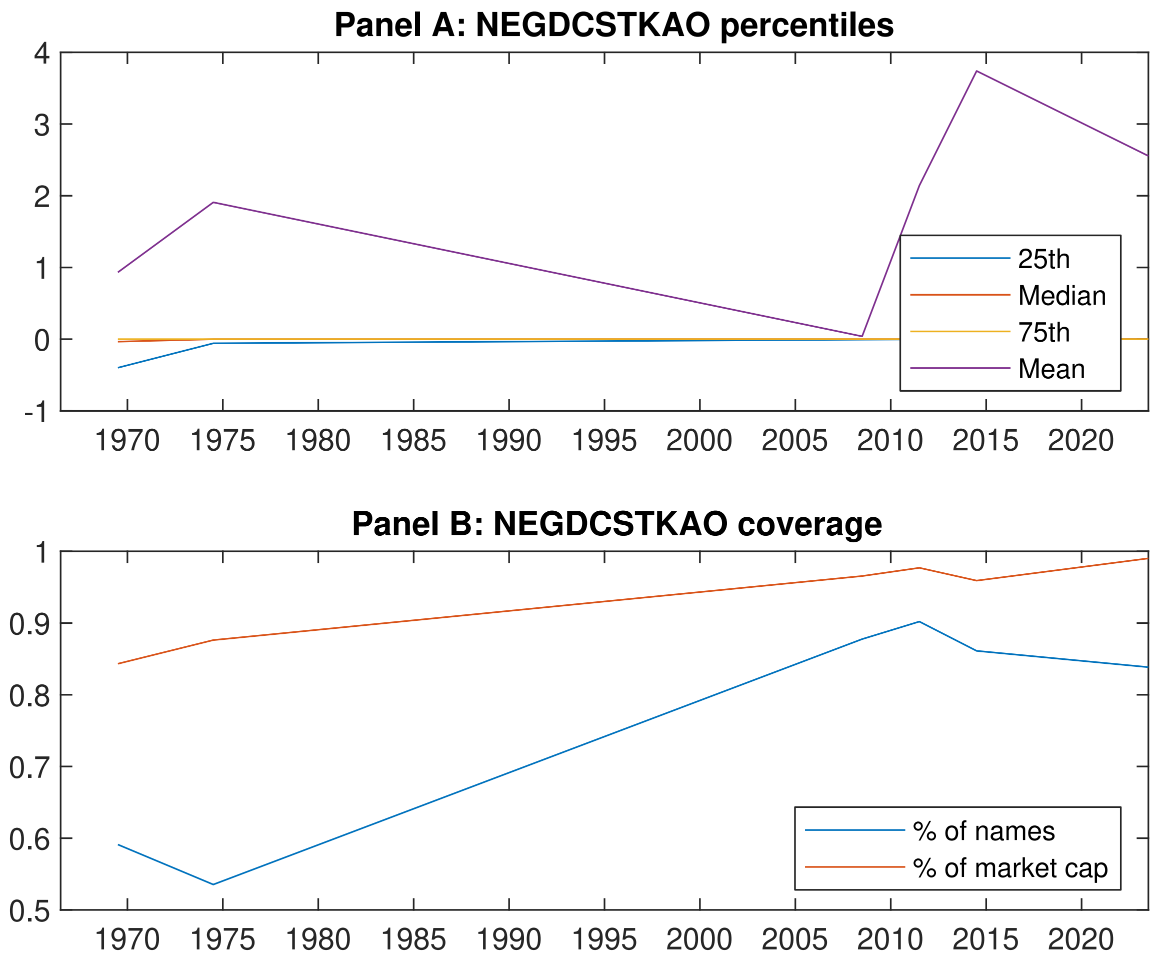

Figure 1: Times series of EDF percentiles and coverage.
This figure plots descriptive statistics for EDF. Panel A shows cross-sectional percentiles of EDF over the sample. Panel B plots the monthly coverage of EDF relative to the universe of CRSP stocks with available market capitalizations.

Table 1: Basic sort: VW, quintile, NYSE-breaks
This table reports average excess returns and alphas for portfolios sorted on EDF. 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 196606 to 202306.

Panel A: Excess returns and alphas on EDF-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.43	0.49	0.59	0.69	0.77	0.34
	[2.46]	[2.59]	[3.12]	[4.08]	[4.58]	[4.61]
α_{CAPM}	-0.12	-0.11	-0.01	0.16	0.24	0.36
	[-2.39]	[-2.45]	[-0.15]	[3.31]	[5.30]	[4.86]
α_{FF3}	-0.11	-0.09	-0.01	0.12	0.20	0.31
	[-2.27]	[-2.12]	[-0.26]	[2.67]	[4.60]	[4.24]
α_{FF4}	-0.10	-0.08	0.03	0.09	0.18	0.27
	[-1.91]	[-1.77]	[0.66]	[1.96]	[4.13]	[3.72]
α_{FF5}	-0.12	-0.05	0.02	0.03	0.10	0.21
	[-2.34]	[-1.06]	[0.43]	[0.65]	[2.35]	[2.94]
α_{FF6}	-0.11	-0.04	0.05	0.01	0.09	0.20
	[-2.08]	[-0.89]	[1.11]	[0.26]	[2.20]	[2.68]
Panel B: Fama and French (2018) 6-factor model loadings for EDF-sorted portfolios						
β_{MKT}	0.96	1.02	1.04	1.00	0.99	0.03
	[80.49]	[94.70]	[91.71]	[95.03]	[100.66]	[1.74]
β_{SMB}	-0.01	0.02	-0.02	-0.07	-0.02	-0.01
	[-0.37]	[1.45]	[-1.17]	[-4.71]	[-1.29]	[-0.48]
β_{HML}	0.02	-0.01	0.02	0.07	0.05	0.03
	[0.94]	[-0.59]	[0.89]	[3.53]	[2.87]	[0.98]
β_{RMW}	0.08	-0.06	-0.04	0.12	0.14	0.07
	[3.24]	[-2.82]	[-1.86]	[5.79]	[7.36]	[1.94]
β_{CMA}	-0.08	-0.10	-0.05	0.19	0.21	0.29
	[-2.51]	[-3.19]	[-1.42]	[6.24]	[7.36]	[5.90]
β_{UMD}	-0.02	-0.01	-0.05	0.03	0.01	0.03
	[-1.61]	[-1.07]	[-4.55]	[2.57]	[0.82]	[1.57]
Panel C: Average number of firms (n) and market capitalization (me)						
n	814	652	551	662	733	
me ($\$10^6$)	1661	1405	2131	2253	2439	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the EDF 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 196606 to 202306.

Panel A: Gross Returns and Alphas								
Portfolios	Breaks	Weights	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Quintile	NYSE	VW	0.34	0.36	0.31	0.27	0.21	0.20
			[4.61]	[4.86]	[4.24]	[3.72]	[2.94]	[2.68]
Quintile	NYSE	EW	0.49	0.57	0.48	0.40	0.32	0.27
			[7.13]	[8.60]	[8.08]	[6.88]	[5.77]	[4.94]
Quintile	Name	VW	0.34	0.35	0.30	0.28	0.24	0.23
			[4.54]	[4.64]	[4.05]	[3.67]	[3.19]	[3.00]
Quintile	Cap	VW	0.31	0.33	0.29	0.25	0.22	0.19
			[4.22]	[4.37]	[3.86]	[3.26]	[2.89]	[2.53]
Decile	NYSE	VW	0.29	0.31	0.24	0.21	0.21	0.18
			[3.27]	[3.41]	[2.74]	[2.31]	[2.27]	[2.00]
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	0.30	0.33	0.28	0.26	0.20	0.20
			[4.11]	[4.41]	[3.88]	[3.63]	[2.83]	[2.72]
Quintile	NYSE	EW	0.30	0.36	0.28	0.24	0.11	0.09
			[3.90]	[4.92]	[4.19]	[3.67]	[1.78]	[1.54]
Quintile	Name	VW	0.31	0.32	0.28	0.27	0.23	0.22
			[4.04]	[4.20]	[3.70]	[3.52]	[3.00]	[2.93]
Quintile	Cap	VW	0.28	0.30	0.26	0.24	0.21	0.19
			[3.75]	[3.96]	[3.52]	[3.21]	[2.79]	[2.60]
Decile	NYSE	VW	0.25	0.27	0.22	0.20	0.18	0.17
			[2.82]	[3.02]	[2.46]	[2.24]	[2.02]	[1.92]

Table 3: Conditional sort on size and EDF

This table presents results for conditional double sorts on size and EDF. 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 EDF. 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 EDF and short stocks with low EDF .Panel B documents the average number of firms and the average firm size for each portfolio. The sample period is 196606 to 202306.

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

Size quintiles	EDF Quintiles					EDF Strategies					
	(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
(1)	0.43	0.70	0.84	0.94	0.95	0.52	0.59	0.51	0.46	0.35	0.32
	[1.61]	[2.69]	[3.24]	[3.72]	[4.02]	[5.97]	[7.00]	[6.61]	[5.90]	[4.68]	[4.27]
(2)	0.54	0.65	0.85	0.91	0.94	0.41	0.47	0.36	0.32	0.27	0.25
	[2.23]	[2.69]	[3.56]	[4.00]	[4.24]	[4.77]	[5.69]	[4.77]	[4.18]	[3.56]	[3.21]
(3)	0.63	0.57	0.79	0.82	0.92	0.29	0.33	0.25	0.24	0.17	0.17
	[2.95]	[2.58]	[3.47]	[3.87]	[4.55]	[3.75]	[4.28]	[3.41]	[3.15]	[2.30]	[2.21]
(4)	0.49	0.59	0.80	0.80	0.83	0.34	0.39	0.30	0.27	0.12	0.11
	[2.40]	[2.86]	[3.84]	[4.02]	[4.40]	[4.14]	[4.75]	[4.02]	[3.52]	[1.65]	[1.48]
(5)	0.43	0.50	0.44	0.55	0.74	0.31	0.31	0.28	0.25	0.26	0.24
	[2.53]	[2.72]	[2.38]	[3.20]	[4.44]	[3.48]	[3.44]	[3.07]	[2.67]	[2.81]	[2.55]

Panel B: Portfolio average number of firms and market capitalization

Size quintiles	EDF Quintiles: Average n					EDF Quintiles: Average market capitalization (\$$10^6$)				
	(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
(1)	370	369	368	366	368	31	34	38	29	29
(2)	109	108	108	108	108	55	56	57	55	56
(3)	79	79	79	78	79	97	96	97	99	100
(4)	66	66	66	66	67	200	205	213	214	216
(5)	61	61	61	61	61	1369	1430	1740	1610	1763

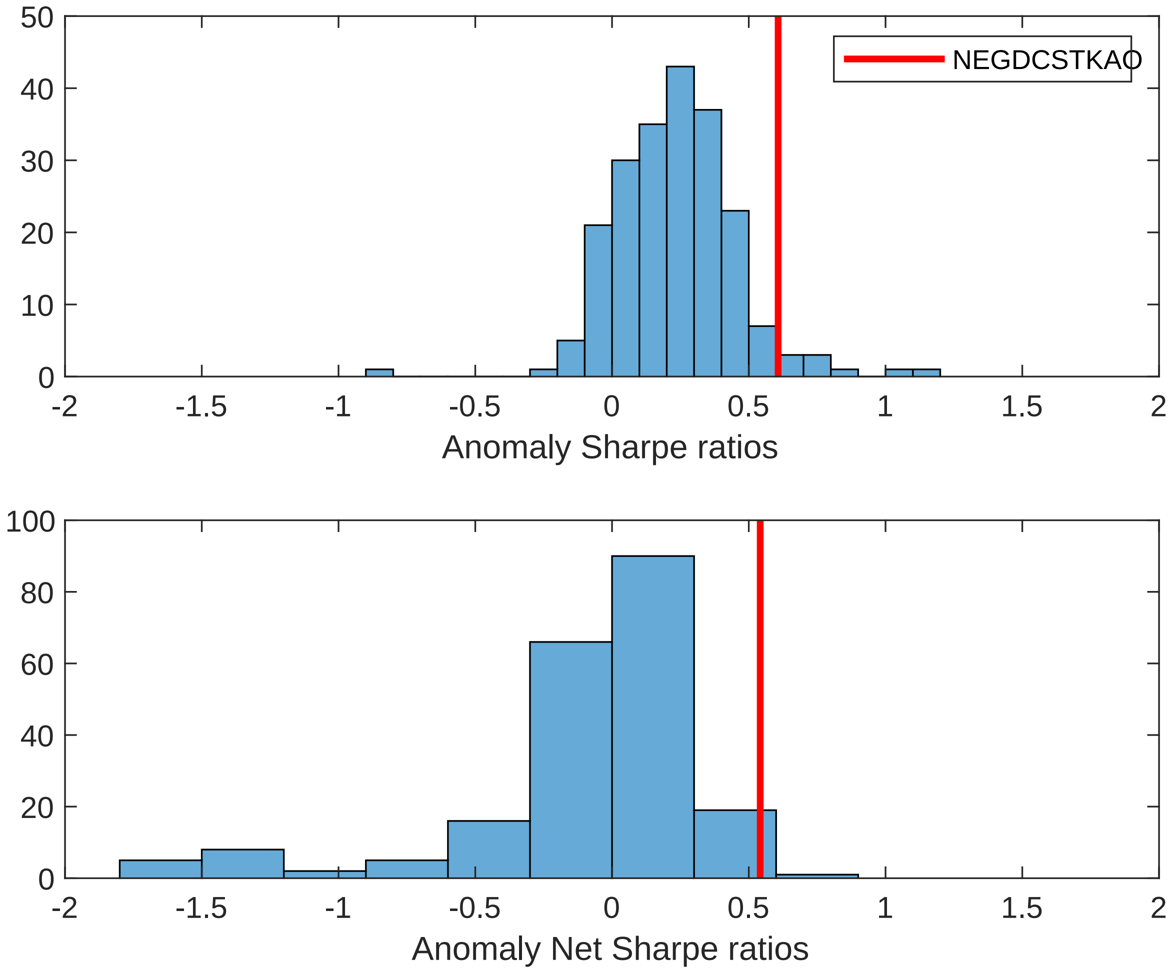

Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the EDF with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

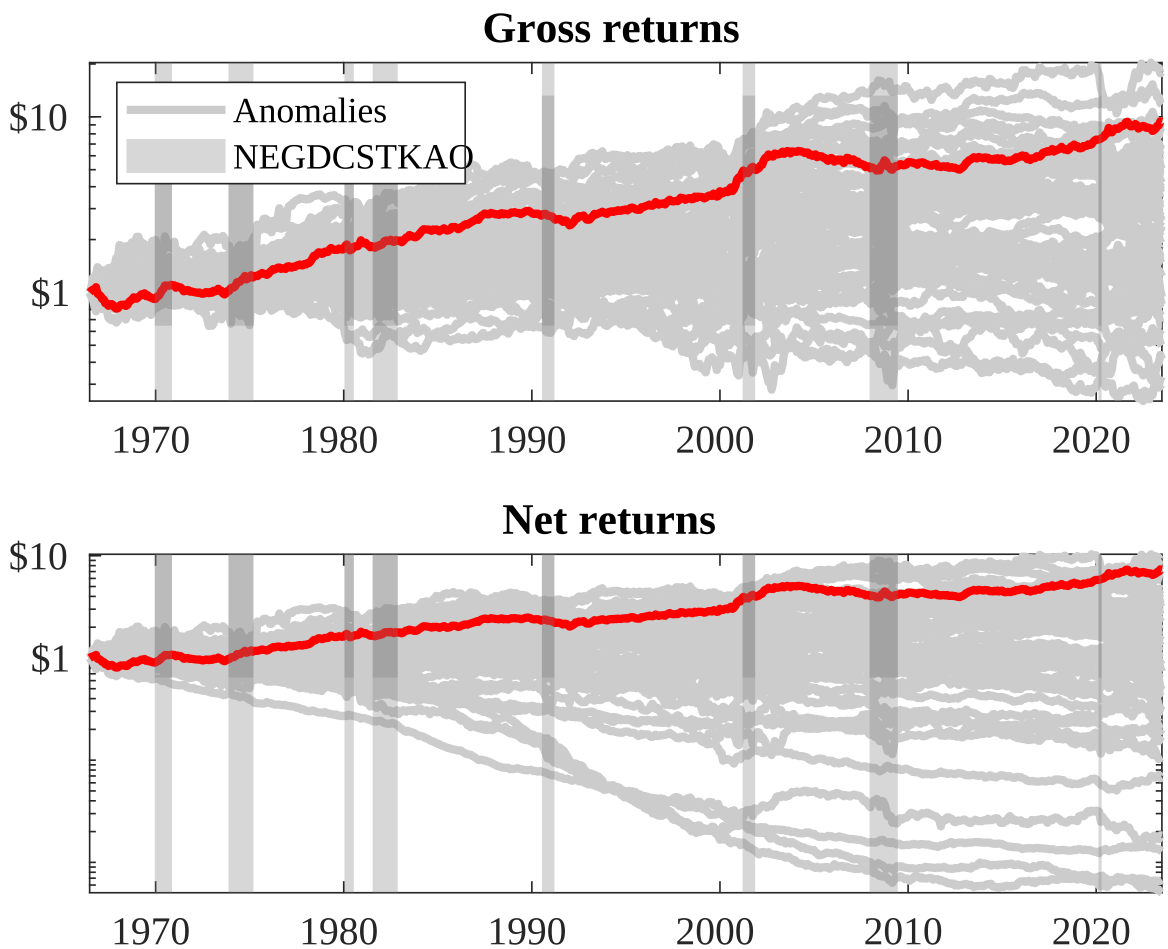

Figure 3: Dollar invested.
This figure plots the growth of a \$1 invested in 212 anomaly trading strategies (gray lines), and compares those with the EDF trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross strategy returns. Panel B plots results for net strtaegy returns.

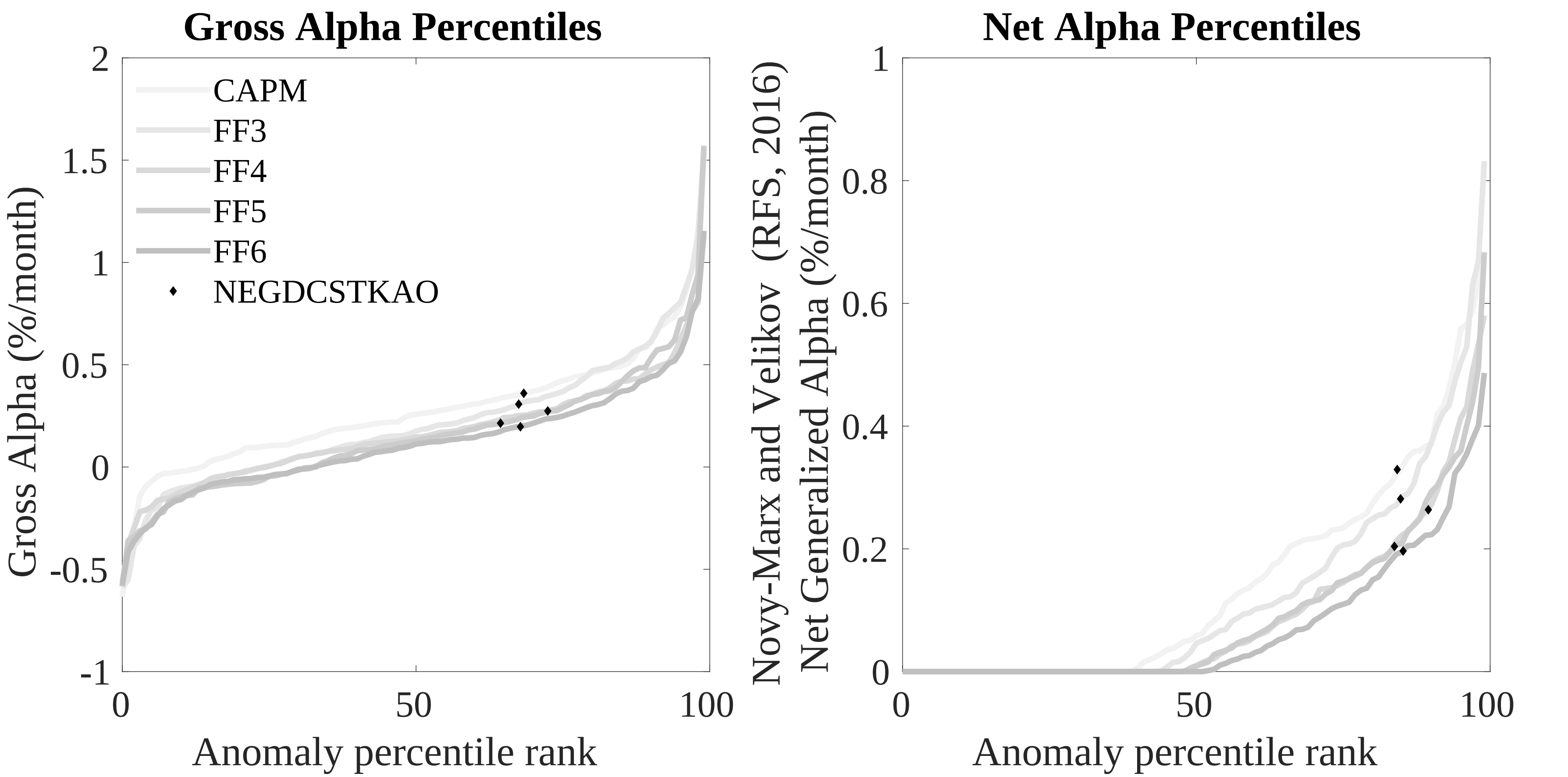

Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the EDF trading strategy alphas (diamonds). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. The alphas include those 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). The left panel plots alphas with no adjustment for trading costs. The right panel plots Novy-Marx and Velikov (2016) net generalized alphas.

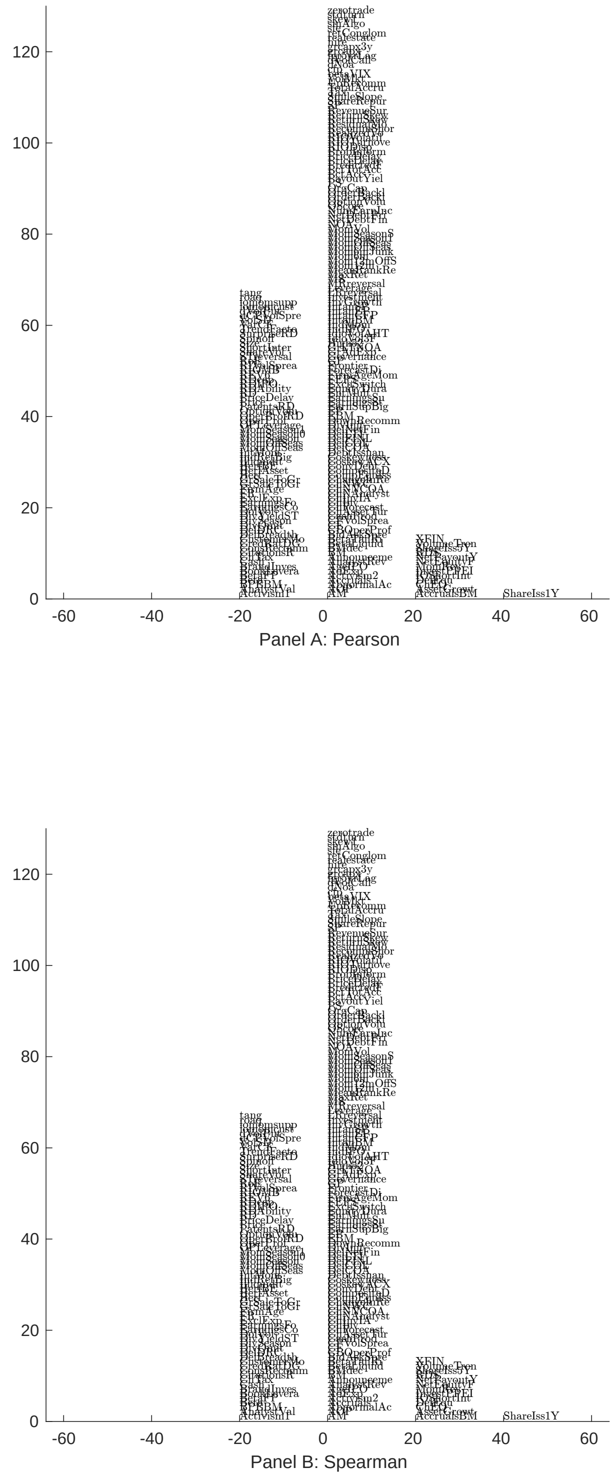

Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 210 filtered anomaly signals with EDF. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

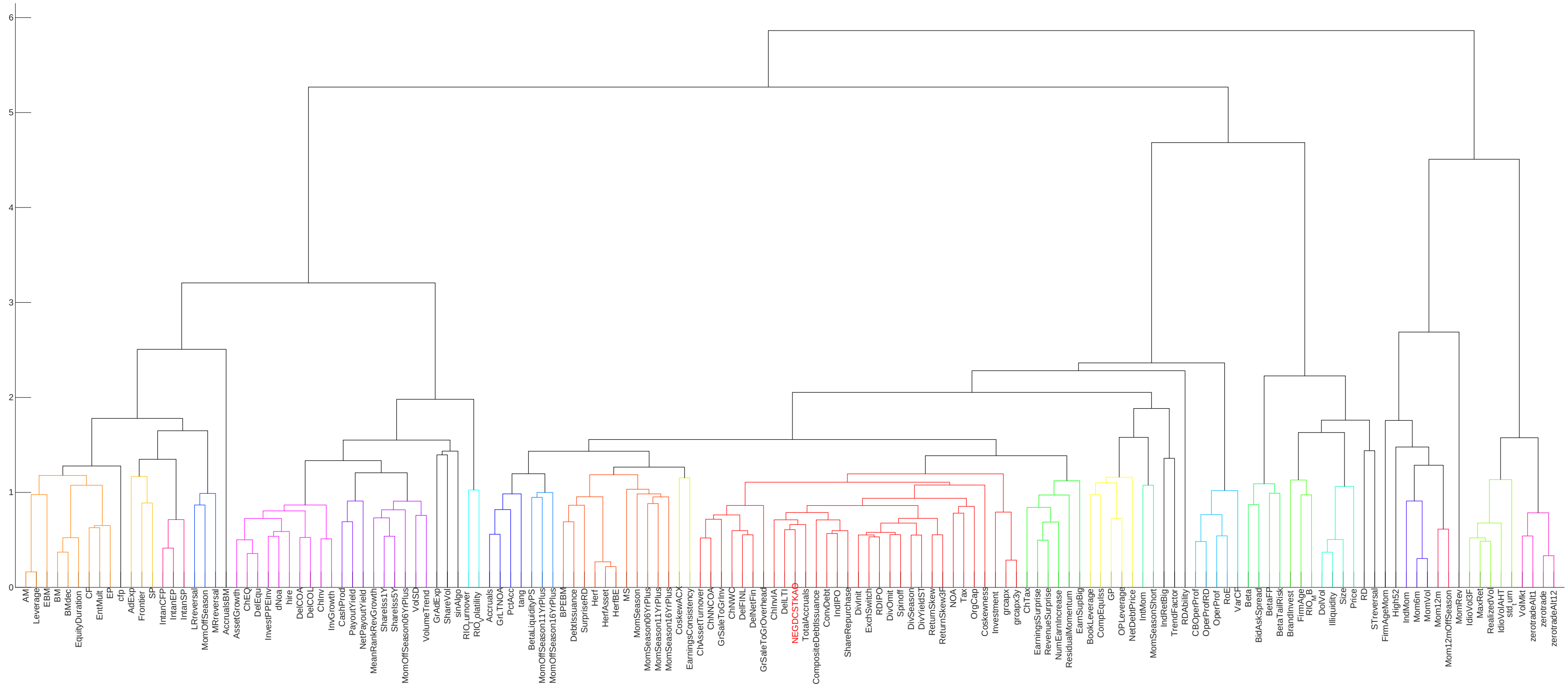

Figure 6: Agglomerative hierarchical cluster plot
This figure plots an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

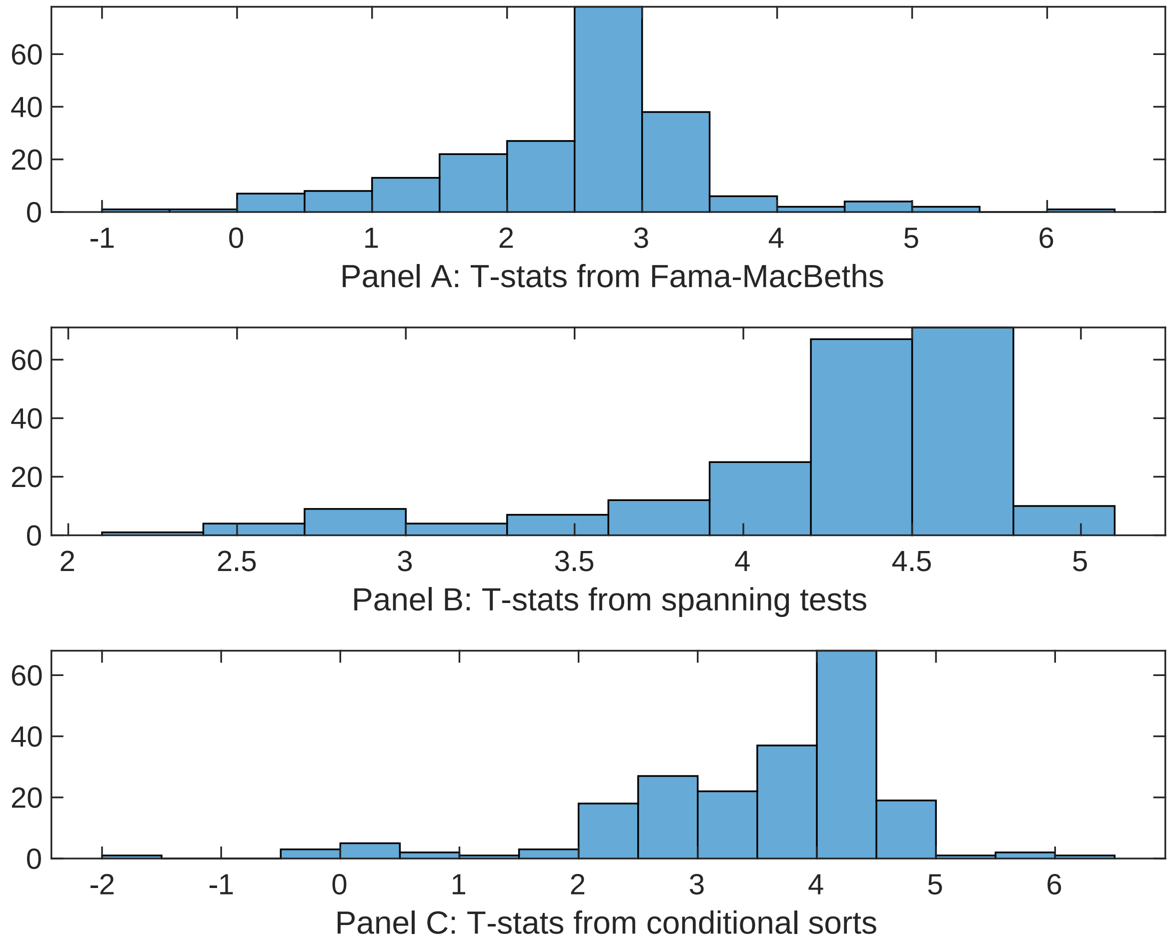

Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of EDF conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{EDF} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{EDF} EDF_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{EDF,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 210 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 210 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on EDF. Stocks are finally grouped into five EDF portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted EDF trading strategies conditioned on each of the 210 filtered anomalies.

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

This table presents Fama-MacBeth results of returns on EDF. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{EDF} EDF_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Share issuance (1 year), Growth in book equity, Net Payout Yield, Share issuance (5 year), Change in equity to assets, Asset growth. 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 196606 to 202306.

Intercept	0.13	0.18	0.12	0.13	0.12	0.13	0.13
	[5.64]	[7.30]	[5.21]	[6.02]	[5.57]	[6.03]	[5.04]
EDF	0.67	0.54	0.30	0.76	0.65	0.54	0.29
	[2.41]	[2.06]	[1.11]	[2.89]	[2.31]	[1.97]	[1.16]
Anomaly 1	0.26						0.88
	[5.97]						[2.23]
Anomaly 2		0.49					-0.41
		[4.52]					[-0.03]
Anomaly 3			0.29				0.24
			[2.54]				[2.22]
Anomaly 4				0.39			0.55
				[4.63]			[0.64]
Anomaly 5					0.15		-0.16
					[4.15]		[-0.27]
Anomaly 6						0.10	0.64
						[8.69]	[5.92]
# months	679	684	679	679	684	684	679
$\bar{R}^2(\%)$	0	0	1	0	0	0	0

Table 5: Spanning tests controlling for most closely related anomalies
This table presents spanning tests results of regressing returns to the EDF trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{EDF} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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 Share issuance (1 year), Growth in book equity, Net Payout Yield, Share issuance (5 year), Change in equity to assets, Asset growth. 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 196606 to 202306.

Intercept	0.17	0.20	0.19	0.17	0.22	0.21	0.16
	[2.46]	[2.79]	[2.66]	[2.33]	[2.99]	[2.79]	[2.39]
Anomaly 1	25.30						17.08
	[6.96]						[4.04]
Anomaly 2		31.97					29.89
		[8.14]					[5.23]
Anomaly 3			14.47				3.48
			[5.17]				[1.09]
Anomaly 4				12.42			-0.59
				[3.27]			[-0.15]
Anomaly 5					20.26		-4.27
					[5.28]		[-0.80]
Anomaly 6						7.35	-11.87
						[1.52]	[-2.35]
mkt	5.22	4.22	5.65	5.03	2.85	3.20	6.12
	[3.12]	[2.53]	[3.27]	[2.86]	[1.67]	[1.84]	[3.57]
smb	0.43	-2.02	2.16	-1.22	-1.27	-1.64	0.94
	[0.18]	[-0.84]	[0.87]	[-0.49]	[-0.51]	[-0.64]	[0.38]
hml	0.83	-0.05	-1.45	0.77	1.13	3.44	-1.96
	[0.25]	[-0.01]	[-0.42]	[0.22]	[0.34]	[1.03]	[-0.57]
rmw	-1.75	7.99	-1.56	4.30	8.33	6.20	0.47
	[-0.50]	[2.46]	[-0.42]	[1.26]	[2.49]	[1.83]	[0.12]
cma	16.99	-3.05	18.71	25.73	7.59	19.75	8.14
	[3.31]	[-0.50]	[3.51]	[5.06]	[1.21]	[2.57]	[1.09]
umd	2.57	2.39	4.11	3.01	3.35	2.95	2.07
	[1.56]	[1.45]	[2.44]	[1.78]	[1.97]	[1.70]	[1.26]
# months	680	684	680	680	684	684	680
$\bar{R}^2$(%)	19	19	16	15	15	11	23

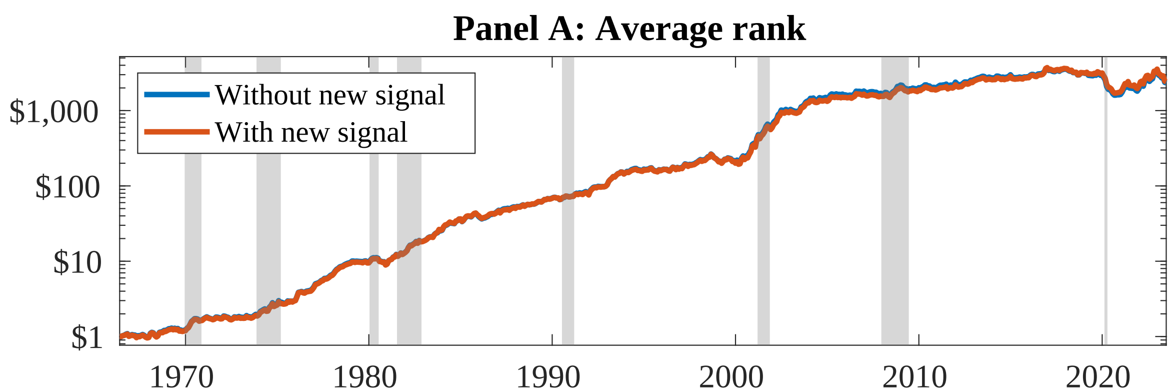

Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 155 anomalies. The red solid lines indicate combination trading strategies that utilize the 155 anomalies as well as EDF. Panel A shows results using "Average rank" as the combination method. See Section 7 for details on the combination methods.

References

Baker, M. and Wurgler, J. (2002). Market timing and capital structure. *The Journal of Finance*, 57(1):1–32.

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.

Daniel, K. and Titman, S. (1999). Market efficiency in an irrational world. *Financial Analysts Journal*, 55(6):28–40.

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

Eckbo, B. E. and Masulis, R. W. (1995). Seasoned equity offerings: A survey. *Finance*, 1:1017–1072.

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.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2020). An augmented q-factor model with expected growth. *Review of Financial Studies*, 34(6):2940–2987.

Loughran, T. and Ritter, J. R. (1995). The new issues puzzle. *The Journal of Finance*, 50(1):23–51.

Myers, S. C. and Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of Financial Economics*, 13(2):187–221.

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*.

Pontiff, J. and Woodgate, A. (2008). Share issuance and cross-sectional returns. *The Journal of Finance*, 63(2):921–945.