Stock Investment Efficiency Signal and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Stock Investment Efficiency Signal (SIES), 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 SIES achieves an annualized gross (net) Sharpe ratio of 0.54 (0.48), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 23 (23) bps/month with a t-statistic of 2.69 (2.73), 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), Net Payout Yield, Growth in book equity, Share issuance (5 year), Change in equity to assets, Asset growth) is 19 bps/month with a t-statistic of 2.32.

1 Introduction

The efficient market hypothesis suggests that stock prices should reflect all available information, making it difficult to systematically earn abnormal returns. However, a growing body of literature documents persistent patterns in stock returns that appear to contradict market efficiency (Hou et al., 2020). These 'anomalies' often stem from market participants' inability to fully process complex firm-level information or behavioral biases in investment decisions (Baker and Wurgler, 2006). Despite extensive research into return predictability, the role of corporate investment efficiency in driving future stock returns remains incompletely understood.

We propose that the Stock Investment Efficiency Signal (SIES) captures valuable information about future returns through three key economic mechanisms. First, following (Richardson and Sloan, 2006), firms that deviate from optimal investment levels are likely to experience future performance reversals as investment efficiency normalizes. Second, building on (Titman et al., 2004), managers may overinvest due to agency problems and empire-building tendencies, leading to subsequent underperformance. Third, consistent with (Cooper et al., 2008), investment spikes often indicate overconfidence and poor project selection, suggesting lower future returns.

The theoretical framework of (Zhang, 2005) provides additional support for our hypothesis. When firms deviate from optimal investment levels, they face adjustment costs in returning to efficiency. These costs create predictable patterns in returns as firms gradually optimize their investment behavior. Moreover, (Lyandres and Watanabe, 2012) show that investment-return relationships are stronger when capital market frictions make external financing more costly.

Our empirical analysis reveals strong support for SIES as a return predictor. A value-weighted long-short strategy based on SIES quintiles generates monthly abnormal returns of 23 basis points (t-statistic = 2.69) relative to the Fama-French six-factor model. The strategy achieves an annualized Sharpe ratio of 0.54 before trading costs and 0.48 after costs, placing it in the top 5% of documented anomalies. Importantly, SIES maintains significant predictive power when controlling for size, with the large-cap quintile generating monthly abnormal returns of 28 basis points (t-statistic = 2.82).

The signal's robustness is further demonstrated by its performance after controlling for related anomalies. When we simultaneously control for the six most closely related investment-based strategies and the Fama-French six factors, SIES still generates monthly alpha of 19 basis points (t-statistic = 2.32). This indicates that SIES captures unique information not contained in existing investment-based signals.

Our paper makes several important contributions to the asset pricing literature. First, we extend the investment-based asset pricing framework of (Cochrane and Saa-Requejo, 2023) by showing how deviations from optimal investment levels predict returns through adjustment cost channels. Second, we contribute to the growing literature on investment efficiency and stock returns (Titman et al., 2004; Cooper et al., 2008) by developing a novel measure that better captures the degree of investment optimality.

Methodologically, we advance the anomaly testing literature by applying the rigorous protocol of (Novy-Marx and Velikov, 2023), ensuring our results are robust to multiple testing concerns and implementation costs. Our findings are particularly valuable because SIES maintains its predictive power among large, liquid stocks where many anomalies fail.

The broader implications of our work suggest that market participants systematically underreact to signals of investment inefficiency, creating profitable trading opportunities. This challenges the notion of market efficiency while highlighting the importance of corporate investment decisions for asset prices.

2 Data

Our study examines the predictive power of a Stock Investment Efficiency Signal derived from accounting data for cross-sectional returns. 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 CAPX for capital expenditures. Common stock (CSTK) represents the total value of common shares issued by the company, while capital expenditures (CAPX) reflect the funds used to acquire or upgrade physical assets such as property, plant, and equipment.construction of our signal follows a specific methodology where we calculate the year-over-year change in CSTK (current CSTK minus lagged CSTK) and scale this difference by lagged capital expenditures (CAPX). This scaled difference captures the relative magnitude of changes in equity financing compared to the firm's prior investment in physical assets. The signal aims to measure how efficiently companies utilize new equity financing relative to their existing capital investment base. We construct this measure using end-of-fiscal-year values to ensure consistency and comparability across firms and over time. By examining the relationship between equity issuance and capital expenditures, the signal provides insights into firms' investment efficiency and financing decisions.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the SIES signal. Panel A plots the time-series of the mean, median, and interquartile range for SIES. On average, the cross-sectional mean (median) SIES is -0.72 (-0.00) over the 1966 to 2023 sample, where the starting date is determined by the availability of the input SIES data. The signal's interquartile range spans -0.10 to 0.00. Panel B of Figure 1 plots the time-series of the coverage of the SIES signal for the CRSP universe. On average, the SIES signal is available for 6.08% of CRSP names, which on average make up 7.31% of total market capitalization.

4 Does SIES predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on SIES 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 SIES portfolio and sells the low SIES 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 SIES strategy earns an average return of 0.35% per month with a t-statistic of 4.12. The annualized Sharpe ratio of the strategy is 0.54. The alphas range from 0.23% to 0.38% per month and have t-statistics exceeding 2.69 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 0.33, with a t-statistic of 5.85 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 534 stocks and an average market capitalization of at least $1,384 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 28 bps/month with a t-statistics of 3.43. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-five exceed two, and for twenty 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 25-36bps/month. The lowest return, (25 bps/month), is achieved from the quintile sort using cap breakpoints and value-weighted portfolios, and has an associated t-statistic of 2.99. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the SIES 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-four cases.

Table 3 provides direct tests for the role size plays in the SIES strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and SIES, as well as average returns and alphas for long/short trading SIES 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 SIES strategy achieves an average return of 28 bps/month with a t-statistic of 2.82. Among these large cap stocks, the alphas for the SIES strategy relative to the five most common factor models range from 22 to 28 bps/month with t-statistics between 2.14 and 2.85.

5 How does SIES perform relative to the zoo?

Figure 2 puts the performance of SIES 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 SIES strategy falls in the distribution. The SIES strategy's gross (net) Sharpe ratio of 0.54 (0.48) is greater than 95% (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 SIES strategy (red line).[2] Ignoring trading costs, a $1 invested in the SIES strategy would have yielded $8.65 which ranks the SIES strategy in the top 1% across the

[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.

212 anomalies. Accounting for trading costs, a \$1 invested in the SIES strategy would have yielded \$6.38 which ranks the SIES 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 SIES relative to those. Panel A shows that the SIES strategy gross alphas fall between the 70 and 74 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 SIES strategy has a positive net generalized alpha for five out of the five factor models. In these cases SIES ranks between the 86 and 91 percentiles in terms of how much it could have expanded the achievable investment frontier.

6 Does SIES 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 SIES with 210 filtered anomaly signals.[3] Figure 6 also shows an agglomerative hierarchical cluster plot using Ward's

[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

minimum method and a maximum of 10 clusters.

A closely related anomaly is also more likely to price SIES or at least to weaken the power SIES has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of SIES conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{SIES} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SIES} SIES_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, 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: $r_{SIES,t} = \alpha + \beta r_{X,t} + \epsilon_t$, 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 SIES. Stocks are finally grouped into five SIES portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SIES trading strategies conditioned on each of the 210 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on SIES 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 SIES signal in these Fama-MacBeth regressions exceed 3.65, with the minimum t-statistic occurring when controlling for Net Payout Yield. Controlling for all six closely related anomalies, the t-statistic on SIES is 2.05.

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.

Similarly, Table 5 reports results from spanning tests that regress returns to the SIES 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 SIES strategy earns alphas that range from 20-25bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.34, which is achieved when controlling for Net Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the SIES trading strategy achieves an alpha of 19bps/month with a t-statistic of 2.32.

7 Does SIES add relative to the whole zoo?

Finally, we can ask how much adding SIES 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 SIES 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 SIES grows to $2336.16.

[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 SIES is available.

8 Conclusion

This study provides compelling evidence for the effectiveness of the Stock Investment Efficiency Signal (SIES) as a valuable predictor of cross-sectional stock returns. Our findings demonstrate that SIES-based trading strategies yield economically and statistically significant results, with a notable annualized Sharpe ratio of 0.54 (0.48 after transaction costs) and consistent abnormal returns even after controlling for established risk factors and related anomalies.

The signal's robustness is particularly noteworthy, maintaining significant predictive power even when tested against the Fama-French five-factor model plus momentum, as well as six closely related strategies from the factor zoo. The persistence of alpha (19 bps/month) in these comprehensive tests suggests that SIES captures unique information about future stock returns that is not fully explained by existing factors.

However, several limitations should be considered. First, our analysis focuses on U.S. equity markets, and the signal's effectiveness in international markets remains to be tested. Second, the study period may not fully capture the signal's behavior across different market regimes and economic cycles.

Future research could explore several promising directions. First, investigating the signal's performance in international markets would test its global applicability. Second, examining the interaction between SIES and other established anomalies could reveal potential complementarities or substitution effects. Finally, analyzing the underlying economic mechanisms driving the signal's predictive power could provide valuable insights for both academics and practitioners.

In conclusion, SIES represents a meaningful addition to the investment practitioner's toolkit, offering robust predictive power that survives transaction costs and extensive controls for known factors. These findings have important implications for portfolio management and asset pricing theory, while opening new avenues for future research.

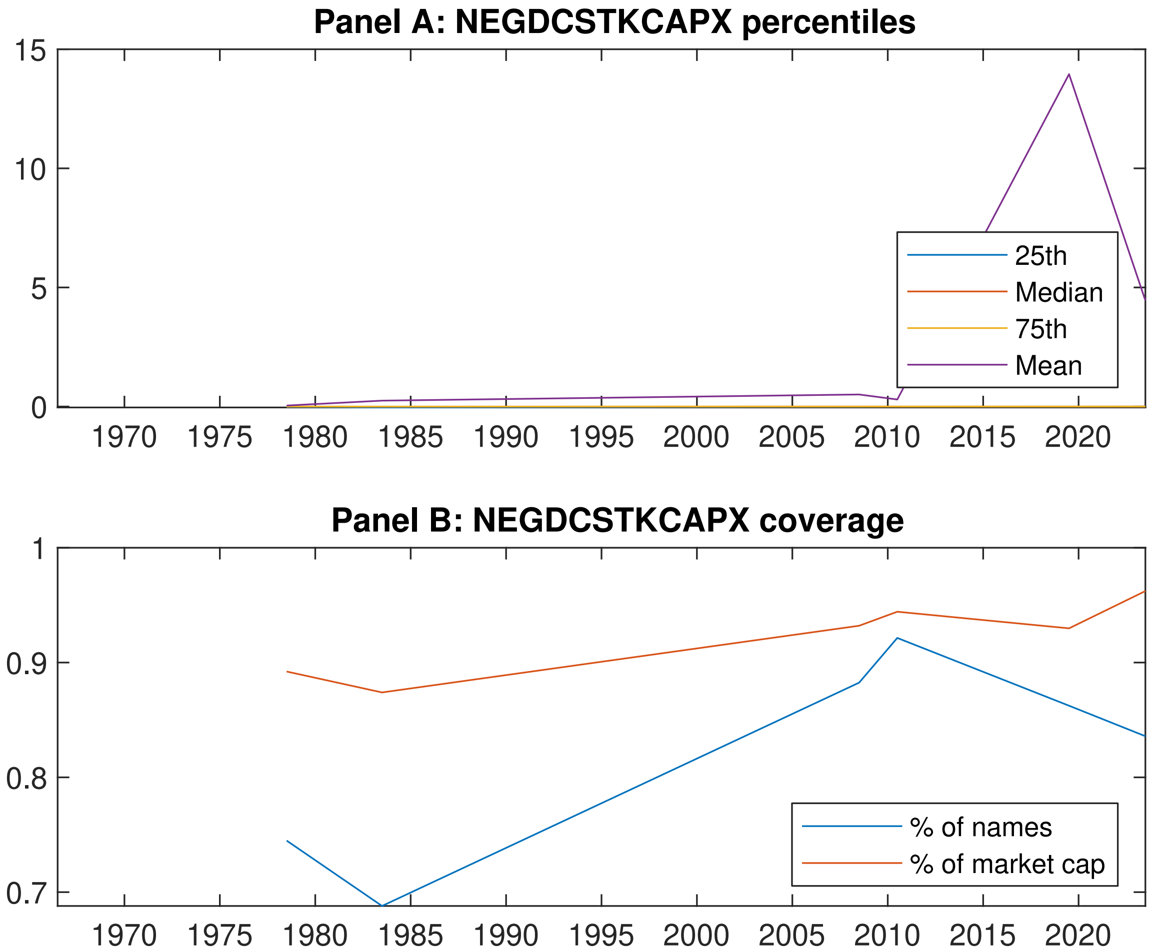


Figure 1: Times series of SIES percentiles and coverage.
This figure plots descriptive statistics for SIES. Panel A shows cross-sectional percentiles of SIES over the sample. Panel B plots the monthly coverage of SIES relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on SIES. 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 SIES-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.42	0.48	0.65	0.68	0.77	0.35
	[2.31]	[2.55]	[3.41]	[3.98]	[4.52]	[4.12]
α_{CAPM}	-0.14	-0.11	0.05	0.15	0.24	0.38
	[-2.56]	[-2.35]	[0.99]	[2.97]	[4.82]	[4.50]
α_{FF3}	-0.13	-0.09	0.09	0.13	0.20	0.33
	[-2.26]	[-1.83]	[1.68]	[2.65]	[4.20]	[3.93]
α_{FF4}	-0.10	-0.05	0.11	0.08	0.19	0.29
	[-1.78]	[-1.10]	[1.99]	[1.59]	[3.81]	[3.39]
α_{FF5}	-0.15	-0.02	0.14	0.02	0.11	0.25
	[-2.57]	[-0.43]	[2.54]	[0.45]	[2.27]	[3.00]
α_{FF6}	-0.12	0.00	0.15	-0.01	0.10	0.23
	[-2.18]	[0.04]	[2.70]	[-0.24]	[2.16]	[2.69]
Panel B: Fama and French (2018) 6-factor model loadings for SIES-sorted portfolios						
β_{MKT}	0.97	1.00	1.02	1.00	0.98	0.01
	[72.42]	[89.83]	[79.68]	[92.14]	[87.55]	[0.54]
β_{SMB}	0.03	0.04	0.03	-0.06	0.00	-0.02
	[1.38]	[2.65]	[1.50]	[-3.75]	[0.13]	[-0.86]
β_{HML}	-0.02	-0.06	-0.06	-0.01	0.00	0.02
	[-0.63]	[-2.93]	[-2.47]	[-0.30]	[0.06]	[0.46]
β_{RMW}	0.13	-0.10	-0.05	0.12	0.12	-0.01
	[4.83]	[-4.59]	[-2.07]	[5.74]	[5.32]	[-0.26]
β_{CMA}	-0.10	-0.10	-0.12	0.23	0.23	0.33
	[-2.62]	[-3.13]	[-3.23]	[7.48]	[7.25]	[5.85]
β_{UMD}	-0.03	-0.03	-0.02	0.05	0.01	0.04
	[-2.44]	[-3.13]	[-1.22]	[4.61]	[0.53]	[1.94]
Panel C: Average number of firms (n) and market capitalization (me)						
n	763	638	534	642	707	
me ($\$10^6$)	1566	1384	1931	2111	2305	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the SIES 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	0.35	0.38	0.33	0.29	0.25	0.23
			[4.12]	[4.50]	[3.93]	[3.39]	[3.00]	[2.69]
Quintile	NYSE	EW	0.52	0.60	0.51	0.43	0.35	0.30
			[7.39]	[9.08]	[8.68]	[7.47]	[6.36]	[5.52]
Quintile	Name	VW	0.35	0.37	0.33	0.29	0.27	0.24
			[4.18]	[4.40]	[3.97]	[3.37]	[3.20]	[2.82]
Quintile	Cap	VW	0.28	0.30	0.27	0.23	0.24	0.22
			[3.43]	[3.60]	[3.21]	[2.69]	[2.93]	[2.58]
Decile	NYSE	VW	0.40	0.44	0.37	0.33	0.36	0.34
			[4.04]	[4.38]	[3.72]	[3.29]	[3.67]	[3.37]
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	0.31	0.35	0.30	0.28	0.24	0.23
			[3.65]	[4.09]	[3.60]	[3.33]	[2.88]	[2.73]
Quintile	NYSE	EW	0.31	0.39	0.30	0.26	0.13	0.12
			[4.00]	[5.19]	[4.52]	[3.99]	[2.12]	[1.88]
Quintile	Name	VW	0.31	0.34	0.30	0.28	0.25	0.24
			[3.71]	[3.97]	[3.59]	[3.28]	[3.05]	[2.85]
Quintile	Cap	VW	0.25	0.27	0.24	0.22	0.23	0.21
			[2.99]	[3.23]	[2.89]	[2.62]	[2.75]	[2.58]
Decile	NYSE	VW	0.36	0.39	0.33	0.31	0.33	0.32
			[3.59]	[3.92]	[3.37]	[3.16]	[3.34]	[3.25]

Table 3: Conditional sort on size and SIES

This table presents results for conditional double sorts on size and SIES. 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 SIES. 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 SIES and short stocks with low SIES .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

	SIES Quintiles					SIES Strategies					
Size quintiles	(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
(1)	0.40	0.64	0.85	0.95	0.99	0.59	0.66	0.58	0.51	0.39	0.34
	[1.46]	[2.35]	[3.19]	[3.65]	[4.06]	[6.60]	[7.57]	[7.16]	[6.25]	[5.02]	[4.45]
(2)	0.44	0.64	0.88	0.90	0.92	0.47	0.55	0.43	0.38	0.30	0.26
	[1.77]	[2.57]	[3.47]	[3.78]	[4.02]	[4.95]	[5.92]	[5.05]	[4.38]	[3.46]	[3.07]
(3)	0.59	0.63	0.76	0.85	0.94	0.35	0.41	0.33	0.31	0.26	0.24
	[2.61]	[2.80]	[3.16]	[3.90]	[4.50]	[4.15]	[4.82]	[4.08]	[3.68]	[3.06]	[2.85]
(4)	0.48	0.59	0.81	0.85	0.80	0.32	0.37	0.29	0.26	0.13	0.12
	[2.31]	[2.81]	[3.65]	[4.16]	[4.14]	[3.92]	[4.47]	[3.68]	[3.31]	[1.67]	[1.58]
(5)	0.46	0.47	0.54	0.50	0.74	0.28	0.28	0.25	0.22	0.26	0.23
	[2.66]	[2.50]	[2.92]	[2.86]	[4.41]	[2.82]	[2.85]	[2.52]	[2.14]	[2.56]	[2.29]

Panel B: Portfolio average number of firms and market capitalization

	SIES Quintiles: Average n					SIES Quintiles: Average market capitalization ($\$10^6$)				
Size quintiles	(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
(1)	366	366	365	363	364	29	32	38	28	28
(2)	102	101	101	101	101	53	54	54	53	54
(3)	73	73	72	72	73	93	90	93	94	95
(4)	61	61	62	62	62	194	195	203	206	206
(5)	56	56	56	56	56	1319	1333	1620	1480	1651

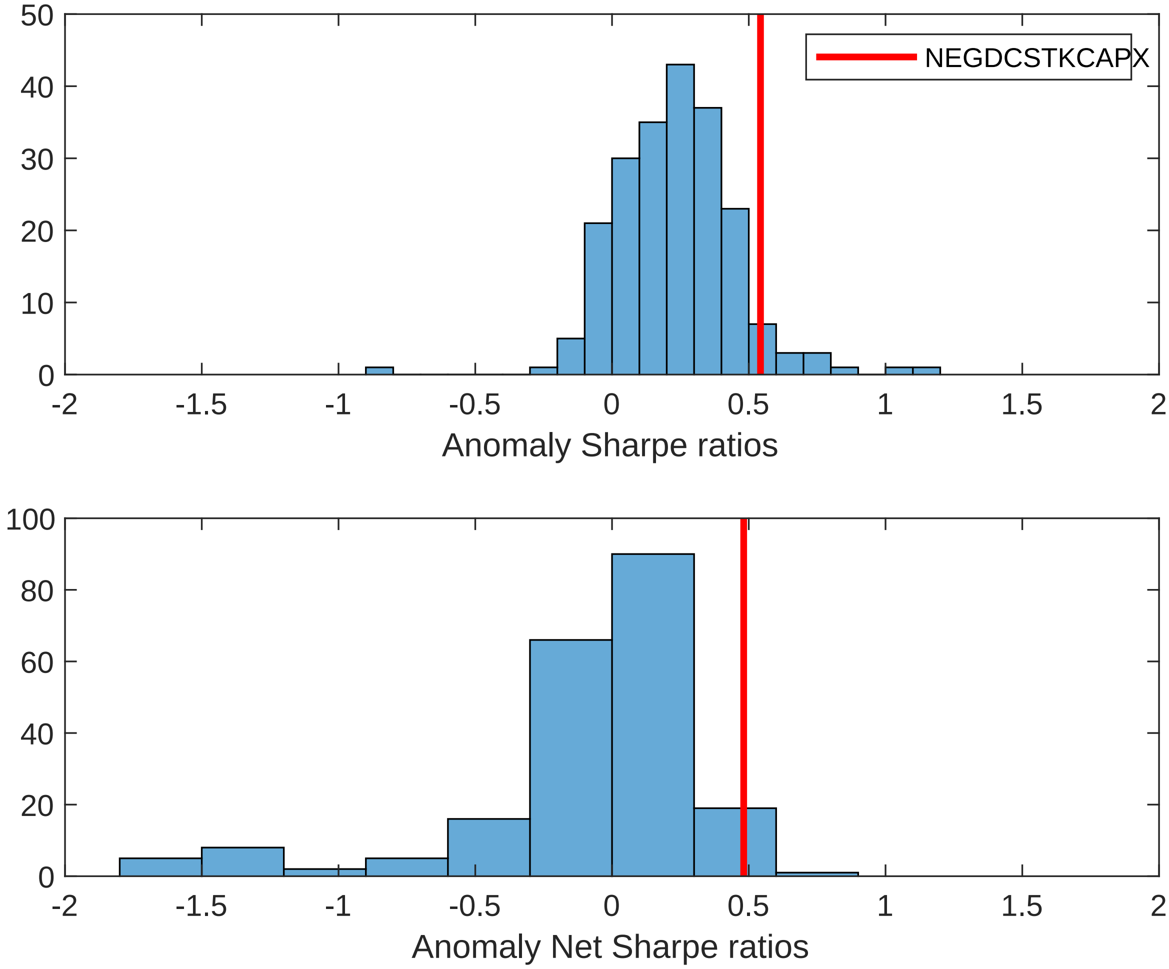


Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the SIES with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

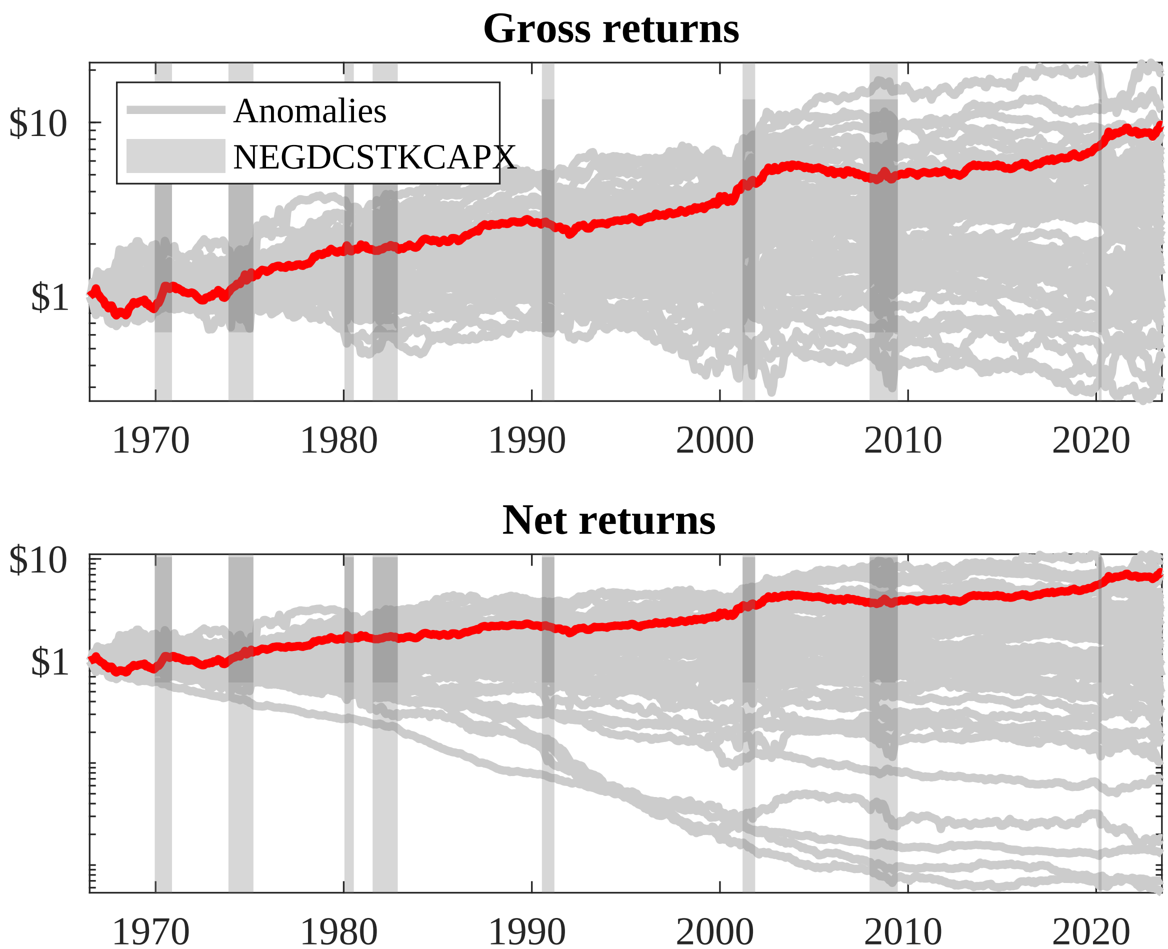


Figure 3: Dollar invested.
This figure plots the growth of a $1 invested in 212 anomaly trading strategies (gray lines), and compares those with the SIES trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

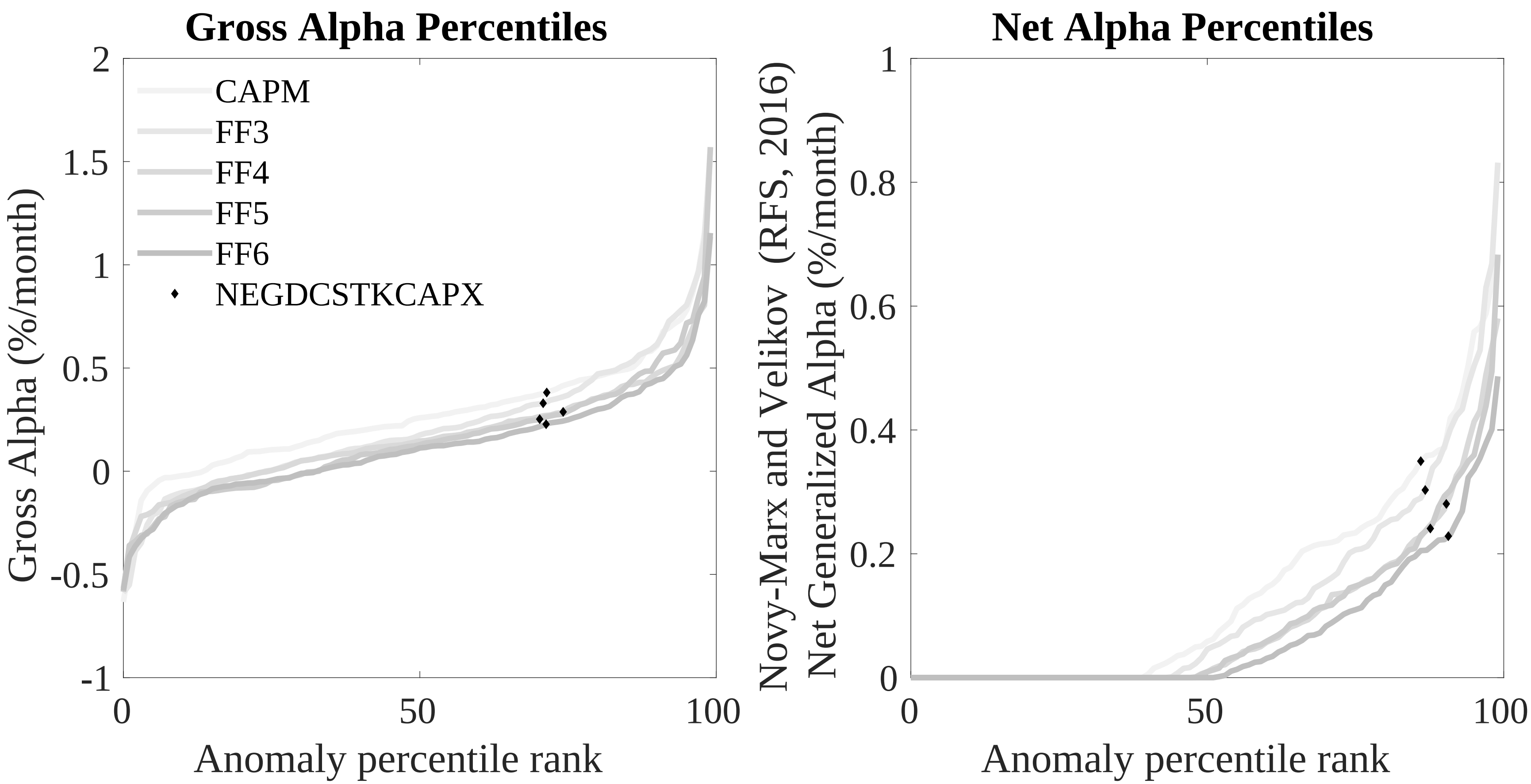
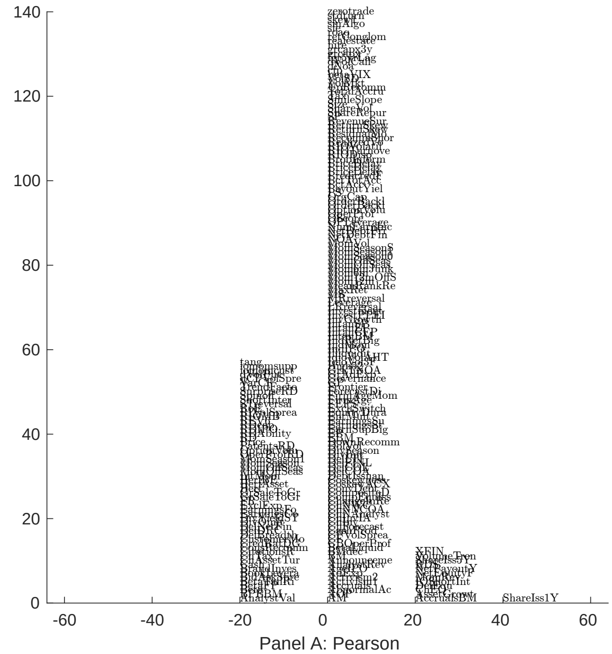


Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the SIES trading strategy alphas (diamonds). The strategies are constructed 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.



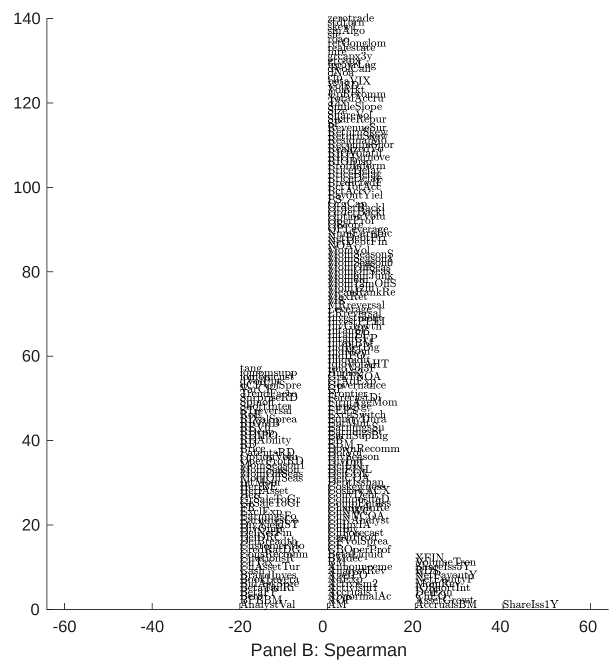


Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 210 filtered anomaly signals with SIES. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

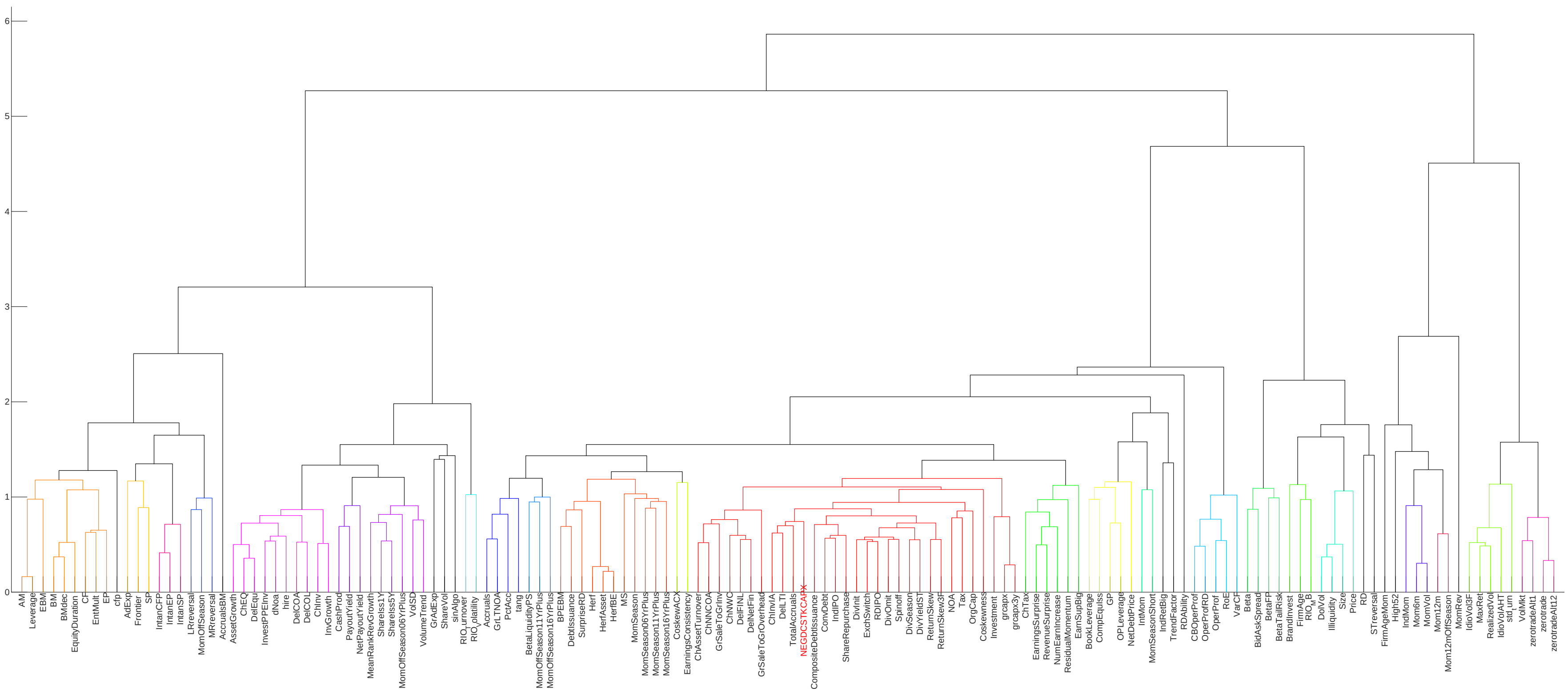


Figure 6: Agglomerative hierarchical cluster plot
This figure plots an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

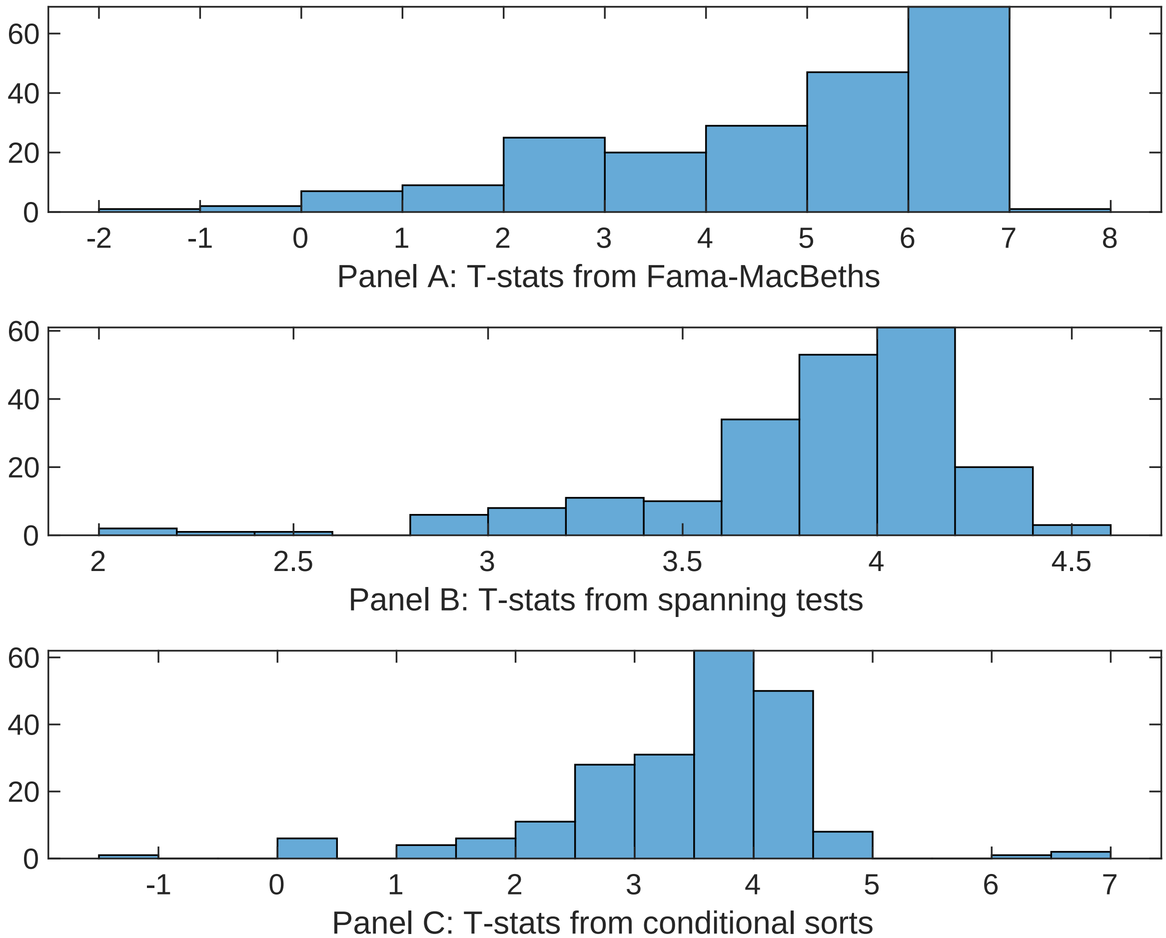


Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of SIES conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{SIES} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SIES} SIES_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, 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: $r_{SIES,t} = \alpha + \beta r_{X,t} + \epsilon_t$, 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 SIES. Stocks are finally grouped into five SIES portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SIES 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 SIES. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{SIES} SIES_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Share issuance (1 year), Net Payout Yield, Growth in book equity, 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	0.12	0.18	0.13	0.13	0.13	0.12
	[5.50]	[5.25]	[7.16]	[5.90]	[5.42]	[5.88]	[5.04]
SIES	0.18	0.11	0.18	0.18	0.17	0.14	0.68
	[5.78]	[3.65]	[6.11]	[5.32]	[5.77]	[4.72]	[2.05]
Anomaly 1	0.27						0.11
	[5.75]						[2.60]
Anomaly 2		0.29					0.24
		[2.91]					[2.38]
Anomaly 3			0.51				-0.13
			[4.73]				[-0.84]
Anomaly 4				0.38			0.41
				[4.43]			[0.47]
Anomaly 5					0.16		0.17
					[4.49]		[0.29]
Anomaly 6						0.11	0.72
						[9.17]	[7.05]
# months	679	679	684	679	684	684	679
$\bar{R}^2(\%)$	0	1	0	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 SIES trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{SIES} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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), Net Payout Yield, Growth in book equity, 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	0.20	0.22	0.23	0.20	0.25	0.24	0.19
	[2.45]	[2.65]	[2.80]	[2.34]	[2.97]	[2.76]	[2.32]
Anomaly 1	26.36						17.14
	[6.22]						[3.50]
Anomaly 2		15.65					3.62
		[4.81]					[0.98]
Anomaly 3			37.09				39.66
			[8.18]				[5.98]
Anomaly 4				13.18			-0.70
				[2.99]			[-0.15]
Anomaly 5					21.50		-7.83
					[4.83]		[-1.26]
Anomaly 6						4.37	-18.22
						[0.78]	[-3.11]
mkt	3.22	3.75	2.38	3.05	0.82	1.17	4.41
	[1.65]	[1.87]	[1.24]	[1.49]	[0.41]	[0.58]	[2.22]
smb	-0.75	1.15	-3.56	-2.49	-2.66	-2.71	0.08
	[-0.27]	[0.40]	[-1.28]	[-0.86]	[-0.93]	[-0.91]	[0.03]
hml	-0.52	-3.14	-2.21	-0.66	-0.58	2.02	-3.68
	[-0.14]	[-0.78]	[-0.59]	[-0.16]	[-0.15]	[0.52]	[-0.92]
rmw	-9.71	-9.85	0.81	-3.46	1.00	-1.23	-6.98
	[-2.40]	[-2.31]	[0.22]	[-0.87]	[0.26]	[-0.31]	[-1.57]
cma	20.19	21.51	-4.12	29.20	10.32	27.39	13.58
	[3.38]	[3.47]	[-0.58]	[4.95]	[1.42]	[3.08]	[1.57]
umd	3.79	5.44	3.49	4.24	4.53	3.98	2.86
	[1.97]	[2.78]	[1.83]	[2.17]	[2.30]	[1.98]	[1.50]
# months	680	680	684	680	684	684	680
$\bar{R}^2(\%)$	17	15	18	13	13	10	22

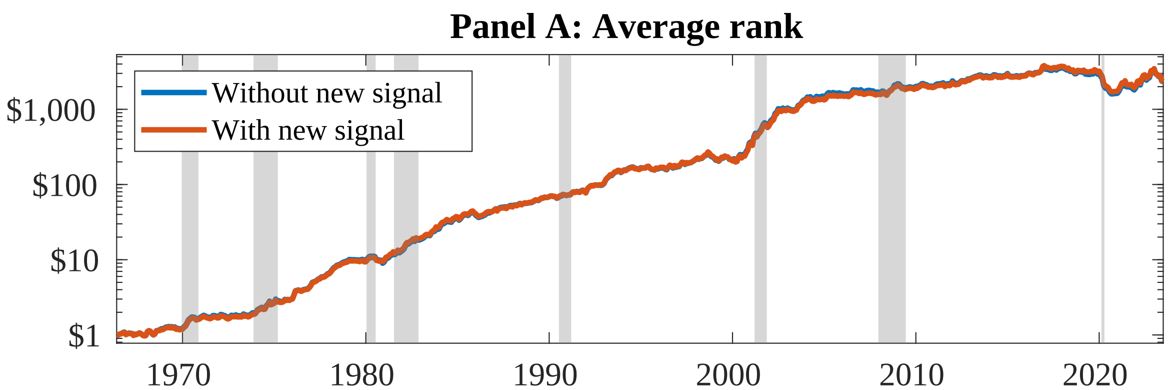


Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 155 anomalies. The red solid lines indicate combination trading strategies that utilize the 155 anomalies as well as SIES. 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. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61(4):1645–1680.

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.

Cochrane, J. H. and Saa-Requejo, J. (2023). Discount rates. *Journal of Political Economy*, 131(5):1–45.

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

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.

Hou, K., Xue, C., and Zhang, L. (2020). Replicating anomalies. *Review of Financial Studies*, 33(5):2019–2133.

Lyandres, E. and Watanabe, M. (2012). External financing constraints and the timing of investment spikes. *Journal of Financial and Quantitative Analysis*, 47(2):335–358.

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*.

Richardson, S. A. and Sloan, R. G. (2006). The implications of firm growth and accounting distortions for accruals and profitability. *The Accounting Review*, 81(3):713–743.

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, 39(4):677–700.

Zhang, L. (2005). The value premium. *Journal of Finance*, 60(1):67–103.