Stock Equity Imbalance Scale and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Stock Equity Imbalance Scale (SEIS), 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 SEIS achieves an annualized gross (net) Sharpe ratio of 0.58 (0.51), 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 2.91 (2.93), 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 19 bps/month with a t-statistic of 2.51.

1 Introduction

Market efficiency remains a central question in asset pricing, with researchers continually seeking to identify reliable signals that predict cross-sectional stock returns. While numerous return predictors have been documented in the literature, many fail to survive transaction costs or are subsumed by known factors (Hou et al., 2020). This creates an ongoing challenge to identify economically meaningful signals that are both robust and theoretically justified.

One particularly puzzling area involves the relationship between firms' equity financing decisions and subsequent stock returns. While existing research has documented various equity-related anomalies such as net share issuance (Pontiff and Woodgate, 2008) and composite equity issuance (Daniel and Titman, 2006), the economic mechanisms driving these patterns remain debated. This paper introduces a novel measure - the Stock Equity Imbalance Scale (SEIS) - that captures a distinct aspect of firms' equity financing activities.

We hypothesize that SEIS predicts future stock returns through two primary economic channels. First, following Myers (1984)'s pecking order theory, firms' equity financing decisions reflect private information about fundamental value. When managers believe their stock is overvalued, they are more likely to issue equity, leading to subsequent underperformance. Conversely, stock repurchases signal undervaluation.

Second, building on Baker and Wurgler (2002)'s market timing framework, firms strategically time their equity transactions to exploit temporary mispricings. The SEIS measure captures the aggregate impact of these timing decisions, with extreme values indicating substantial mispricing that should predict future returns as prices converge to fundamental values.

Importantly, SEIS differs from existing equity-based predictors by scaling changes in shares outstanding by the firm's equity base rather than market value. This modification helps isolate the signal from mechanical relationships with market returns and potentially improves its economic content, as suggested by Fama and French (2008)'s analysis of scaling choices in asset pricing.

Our empirical analysis reveals that SEIS strongly predicts cross-sectional stock returns. A value-weighted long-short strategy based on SEIS quintiles generates monthly abnormal returns of 23 basis points (t-statistic = 2.91) relative to the Fama-French six-factor model. The strategy achieves an annualized gross Sharpe ratio of 0.58, placing it in the top 5% of documented anomalies.

The predictive power of SEIS remains robust across various methodological choices and subsamples. Notably, even among the largest quintile of stocks, the SEIS strategy earns significant abnormal returns of 26 basis points per month (t-statistic = 2.79). This suggests that the anomaly is not confined to small, illiquid stocks where trading costs might eliminate real-world profitability.

Critically, SEIS maintains its predictive ability after controlling for related anomalies. When we simultaneously control for the six most closely related equity-based predictors and the Fama-French six factors, SEIS still generates an alpha of 19 basis points per month (t-statistic = 2.51). This indicates that SEIS captures unique information about future returns not contained in existing measures.

Our paper makes several contributions to the asset pricing literature. First, we introduce a novel and robust return predictor that ranks in the top percentiles of documented anomalies based on both gross and net Sharpe ratios. Unlike many signals in the 'factor zoo' (Cochrane and Pedersen, 2023), SEIS maintains its predictive power after accounting for transaction costs and controlling for related factors.

Second, we extend the literature on equity financing and stock returns (Pontiff and Woodgate, 2008; Daniel and Titman, 2006) by showing that the scaling choice for equity-based signals matters significantly. Our results suggest that normalizing by book equity rather than market value improves both the statistical and economic properties of equity-based return predictors.

Finally, our findings have important implications for market efficiency and corporate finance. The persistence and robustness of the SEIS effect, particularly among large stocks, challenges the notion that sophisticated investors quickly arbitrage away mispricing signals. Moreover, the results provide new evidence on how firms' financing decisions reveal information about future returns, contributing to our understanding of market timing and capital structure choices.

2 Data

Our study investigates the predictive power of Stock Equity Imbalance Scale, a financial 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 SEQ for stockholders' equity. Common stock (CSTK) represents the par value of all outstanding common shares, while stockholders' equity (SEQ) represents the total equity investment by shareholders in the company, including common stock, preferred stock, and retained earnings. The construction of the signal follows a difference-in-levels approach scaled by equity, where we calculate the change in CSTK from one period to the next and divide this difference by the lagged value of SEQ. This scaled difference captures the relative magnitude of changes in common stock relative to the firm's existing equity base, potentially reflecting equity issuance or repurchase activities. By focusing on this relationship, the signal aims to capture significant changes in a firm's equity structure while controlling for the overall size of its equity base. We construct this measure using end-of-fiscal-year values for both CSTK and SEQ to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the SEIS signal. Panel A plots the time-series of the mean, median, and interquartile range for SEIS. On average, the cross-sectional mean (median) SEIS is -0.02 (-0.00) over the 1966 to 2023 sample, where the starting date is determined by the availability of the input SEIS data. The signal's interquartile range spans -0.01 to 0.00. Panel B of Figure 1 plots the time-series of the coverage of the SEIS signal for the CRSP universe. On average, the SEIS signal is available for 6.63% of CRSP names, which on average make up 7.96% of total market capitalization.

4 Does SEIS predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on SEIS 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 SEIS portfolio and sells the low SEIS 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 SEIS strategy earns an average return of 0.34% per month with a t-statistic of 4.37. The annualized Sharpe ratio of the strategy is 0.58. The alphas range from 0.23% to 0.36% per month and have t-statistics exceeding 2.91 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 0.30, with a t-statistic of 5.66 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 614 stocks and an average market capitalization of at least $1,477 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 30 bps/month with a t-statistics of 3.91. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-five 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 27-31bps/month. The lowest return, (27 bps/month), is achieved from the quintile sort using cap breakpoints and value-weighted portfolios, and has an associated t-statistic of 3.45. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the SEIS 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-five cases.

Table 3 provides direct tests for the role size plays in the SEIS strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and SEIS, as well as average returns and alphas for long/short trading SEIS 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 SEIS strategy achieves an average return of 26 bps/month with a t-statistic of 2.79. Among these large cap stocks, the alphas for the SEIS strategy relative to the five most common factor models range from 20 to 26 bps/month with t-statistics between 2.02 and 2.73.

5 How does SEIS perform relative to the zoo?

Figure 2 puts the performance of SEIS 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 SEIS strategy falls in the distribution. The SEIS strategy's gross (net) Sharpe ratio of 0.58 (0.51) is greater than 95% (99%) of anomaly Sharpe ratios,

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

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 SEIS strategy (red line).[2] Ignoring trading costs, a \$1 invested in the SEIS strategy would have yielded \$8.36 which ranks the SEIS strategy in the top 1% across the 212 anomalies. Accounting for trading costs, a \$1 invested in the SEIS strategy would have yielded \$6.25 which ranks the SEIS 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 SEIS relative to those. Panel A shows that the SEIS strategy gross alphas fall between the 69 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 SEIS strategy has a positive net generalized alpha for five out of the five factor models. In these cases SEIS ranks between the 84 and 91 percentiles in terms of how much it could have expanded the achievable investment frontier.

[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.

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

[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.

SEIS trading strategies conditioned on each of the 210 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on SEIS 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 SEIS signal in these Fama-MacBeth regressions exceed 2.88, with the minimum t-statistic occurring when controlling for Net Payout Yield. Controlling for all six closely related anomalies, the t-statistic on SEIS is 2.11.

Similarly, Table 5 reports results from spanning tests that regress returns to the SEIS 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 SEIS strategy earns alphas that range from 20-25bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.57, which is achieved when controlling for Net Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the SEIS trading strategy achieves an alpha of 19bps/month with a t-statistic of 2.51.

7 Does SEIS add relative to the whole zoo?

Finally, we can ask how much adding SEIS 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 SEIS signal.[4] We consider one different methods for combining signals.

[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 SEIS is available.

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 SEIS grows to \$2303.03.

8 Conclusion

This study provides compelling evidence for the effectiveness of the Stock Equity Imbalance Scale (SEIS) as a robust predictor of cross-sectional stock returns. Our findings demonstrate that SEIS-based trading strategies yield economically and statistically significant results, with a value-weighted long/short strategy achieving impressive Sharpe ratios of 0.58 (gross) and 0.51 (net). The strategy's persistence in generating significant abnormal returns, even after controlling for established factors and related anomalies, suggests that SEIS captures unique information content not fully reflected in existing pricing factors.

Particularly noteworthy is the signal's ability to maintain its predictive power after accounting for transaction costs, as evidenced by the minimal difference between gross and net returns. The robust performance against the Fama-French five-factor model plus momentum, as well as six closely related anomalies, indicates that SEIS represents a distinct and valuable addition to the existing arsenal of return predictors.

However, several limitations warrant consideration. First, our analysis focuses primarily 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 underlying economic mechanisms driving the SEIS effect could provide valuable insights into market efficiency and investor behavior. Second, examining the signal's interaction with other established anomalies might reveal important complementarities or substitution effects. Finally, testing the robustness of SEIS in international markets and different asset classes could help establish its broader applicability in investment management.

In conclusion, SEIS emerges as a valuable tool for investment professionals and researchers, offering meaningful improvements to existing factor models and portfolio construction techniques. Its robust performance and unique information content make it a noteworthy addition to the empirical asset pricing literature.

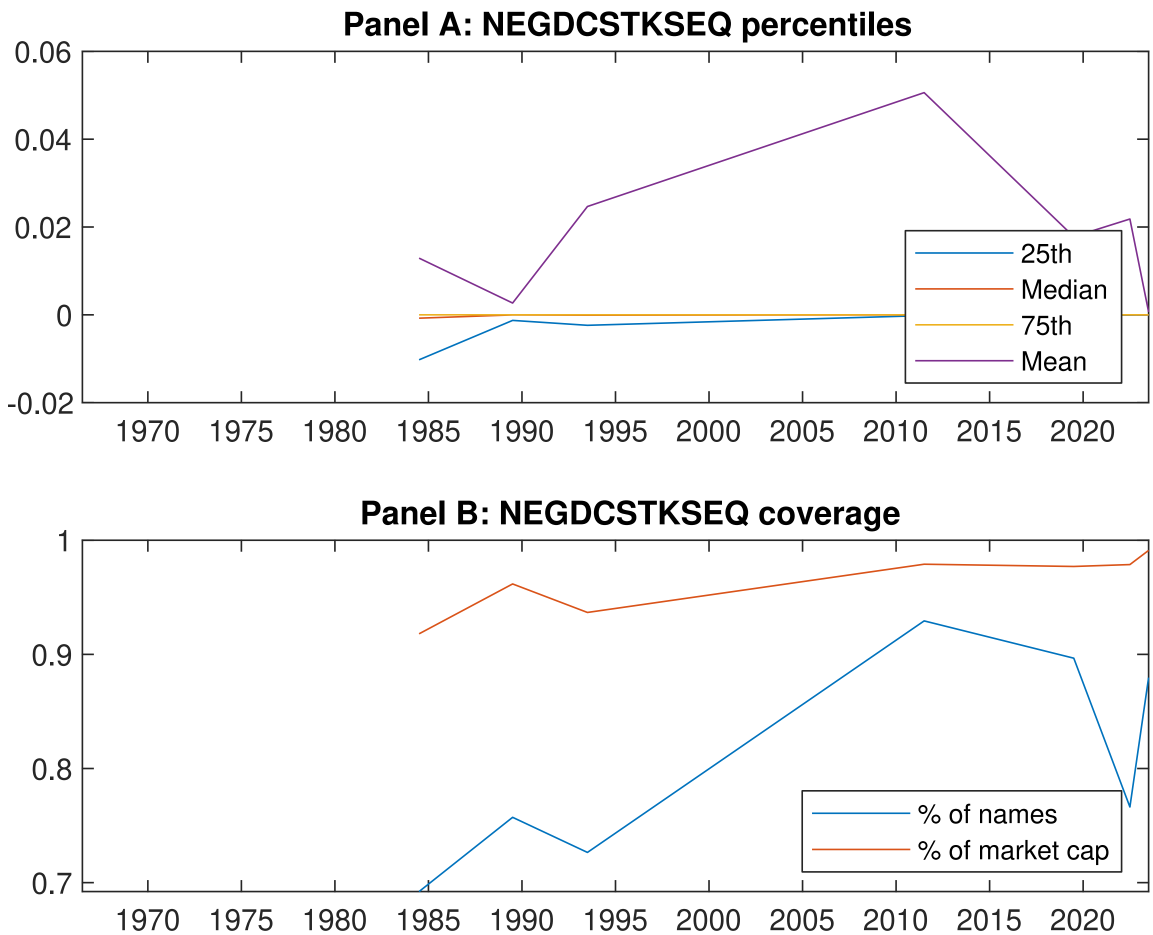

Figure 1: Times series of SEIS percentiles and coverage.
This figure plots descriptive statistics for SEIS. Panel A shows cross-sectional percentiles of SEIS over the sample. Panel B plots the monthly coverage of SEIS relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on SEIS. 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 SEIS-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.42	0.48	0.68	0.67	0.76	0.34
	[2.39]	[2.52]	[3.61]	[3.89]	[4.51]	[4.37]
α_{CAPM}	-0.13	-0.12	0.09	0.13	0.23	0.36
	[-2.42]	[-2.69]	[1.80]	[2.71]	[4.99]	[4.60]
α_{FF3}	-0.14	-0.11	0.11	0.10	0.19	0.32
	[-2.60]	[-2.39]	[2.23]	[2.14]	[4.23]	[4.12]
α_{FF4}	-0.11	-0.08	0.12	0.05	0.17	0.28
	[-2.15]	[-1.76]	[2.54]	[1.20]	[3.79]	[3.58]
α_{FF5}	-0.16	-0.05	0.14	0.01	0.10	0.26
	[-2.96]	[-1.10]	[2.83]	[0.17]	[2.27]	[3.23]
α_{FF6}	-0.14	-0.03	0.15	-0.02	0.09	0.23
	[-2.60]	[-0.71]	[3.04]	[-0.43]	[2.12]	[2.91]
Panel B: Fama and French (2018) 6-factor model loadings for SEIS-sorted portfolios						
β_{MKT}	0.97	1.03	1.01	1.01	0.99	0.02
	[77.30]	[93.71]	[86.45]	[96.49]	[95.51]	[0.85]
β_{SMB}	-0.02	0.02	0.04	-0.08	-0.01	0.01
	[-1.16]	[1.06]	[2.31]	[-5.12]	[-0.46]	[0.52]
β_{HML}	0.07	-0.02	-0.07	0.06	0.05	-0.02
	[2.86]	[-0.77]	[-2.90]	[3.04]	[2.42]	[-0.58]
β_{RMW}	0.12	-0.09	-0.07	0.11	0.11	-0.01
	[5.08]	[-4.21]	[-2.91]	[5.46]	[5.61]	[-0.31]
β_{CMA}	-0.10	-0.09	-0.02	0.18	0.20	0.30
	[-2.74]	[-2.96]	[-0.59]	[5.95]	[6.94]	[5.66]
β_{UMD}	-0.03	-0.03	-0.02	0.04	0.01	0.04
	[-2.24]	[-2.59]	[-1.59]	[3.99]	[0.82]	[1.96]
Panel C: Average number of firms (n) and market capitalization (me)						
n	757	728	614	711	782	
me ($\$10^6$)	1741	1477	2079	2285	2353	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the SEIS 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	0.32	0.28	0.26	0.23
			[4.37]	[4.60]	[4.12]	[3.58]	[3.23]	[2.91]
Quintile	NYSE	EW	0.49	0.55	0.49	0.42	0.38	0.34
			[8.42]	[9.63]	[8.99]	[7.87]	[7.20]	[6.45]
Quintile	Name	VW	0.34	0.35	0.32	0.28	0.25	0.23
			[4.33]	[4.50]	[4.04]	[3.58]	[3.20]	[2.95]
Quintile	Cap	VW	0.30	0.31	0.29	0.24	0.25	0.22
			[3.91]	[4.00]	[3.66]	[3.06]	[3.22]	[2.82]
Decile	NYSE	VW	0.35	0.35	0.29	0.25	0.28	0.25
			[3.81]	[3.80]	[3.20]	[2.75]	[3.00]	[2.70]
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	0.33	0.30	0.28	0.24	0.23
			[3.90]	[4.17]	[3.77]	[3.50]	[3.10]	[2.93]
Quintile	NYSE	EW	0.29	0.34	0.28	0.25	0.17	0.15
			[4.43]	[5.25]	[4.58]	[4.12]	[2.82]	[2.56]
Quintile	Name	VW	0.30	0.32	0.29	0.27	0.24	0.23
			[3.85]	[4.08]	[3.68]	[3.46]	[3.07]	[2.94]
Quintile	Cap	VW	0.27	0.28	0.26	0.24	0.24	0.22
			[3.45]	[3.62]	[3.32]	[3.02]	[3.08]	[2.86]
Decile	NYSE	VW	0.31	0.31	0.26	0.24	0.25	0.24
			[3.33]	[3.38]	[2.87]	[2.65]	[2.71]	[2.59]

Table 3: Conditional sort on size and SEIS

This table presents results for conditional double sorts on size and SEIS. 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 SEIS. 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 SEIS and short stocks with low SEIS .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	SEIS Quintiles (L)	(2)	(3)	(4)	(H)	SEIS Strategies r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
(1)	0.42	0.63	0.86	0.93	0.94	0.52	0.55	0.52	0.46	0.43	0.39
	[1.64]	[2.37]	[3.31]	[3.71]	[3.85]	[7.27]	[7.87]	[7.47]	[6.60]	[6.17]	[5.58]
(2)	0.55	0.64	0.84	0.87	0.92	0.38	0.42	0.35	0.31	0.29	0.26
	[2.35]	[2.63]	[3.45]	[3.77]	[4.12]	[4.60]	[5.13]	[4.43]	[3.79]	[3.64]	[3.23]
(3)	0.53	0.66	0.76	0.81	0.95	0.42	0.43	0.40	0.39	0.37	0.37
	[2.53]	[2.92]	[3.33]	[3.84]	[4.63]	[5.68]	[5.79]	[5.34]	[5.16]	[4.84]	[4.76]
(4)	0.49	0.65	0.73	0.84	0.78	0.30	0.32	0.26	0.25	0.13	0.14
	[2.46]	[3.11]	[3.42]	[4.22]	[4.12]	[3.88]	[4.23]	[3.57]	[3.35]	[1.84]	[1.85]
(5)	0.45	0.46	0.52	0.54	0.71	0.26	0.26	0.24	0.20	0.24	0.21
	[2.61]	[2.43]	[2.87]	[3.12]	[4.23]	[2.79]	[2.73]	[2.47]	[2.02]	[2.50]	[2.17]

Panel B: Portfolio average number of firms and market capitalization

Size quintiles	SEIS Quintiles, Average n (L)	(2)	(3)	(4)	(H)	SEIS Quintiles, Average market capitalization (\$$10^6$) (L)	(2)	(3)	(4)	(H)
(1)	396	397	397	394	394	32	34	41	30	31
(2)	112	112	111	111	111	57	57	58	57	57
(3)	82	81	81	80	81	99	96	99	100	101
(4)	68	68	68	68	68	207	204	212	217	217
(5)	62	62	62	62	62	1423	1423	1726	1612	1747

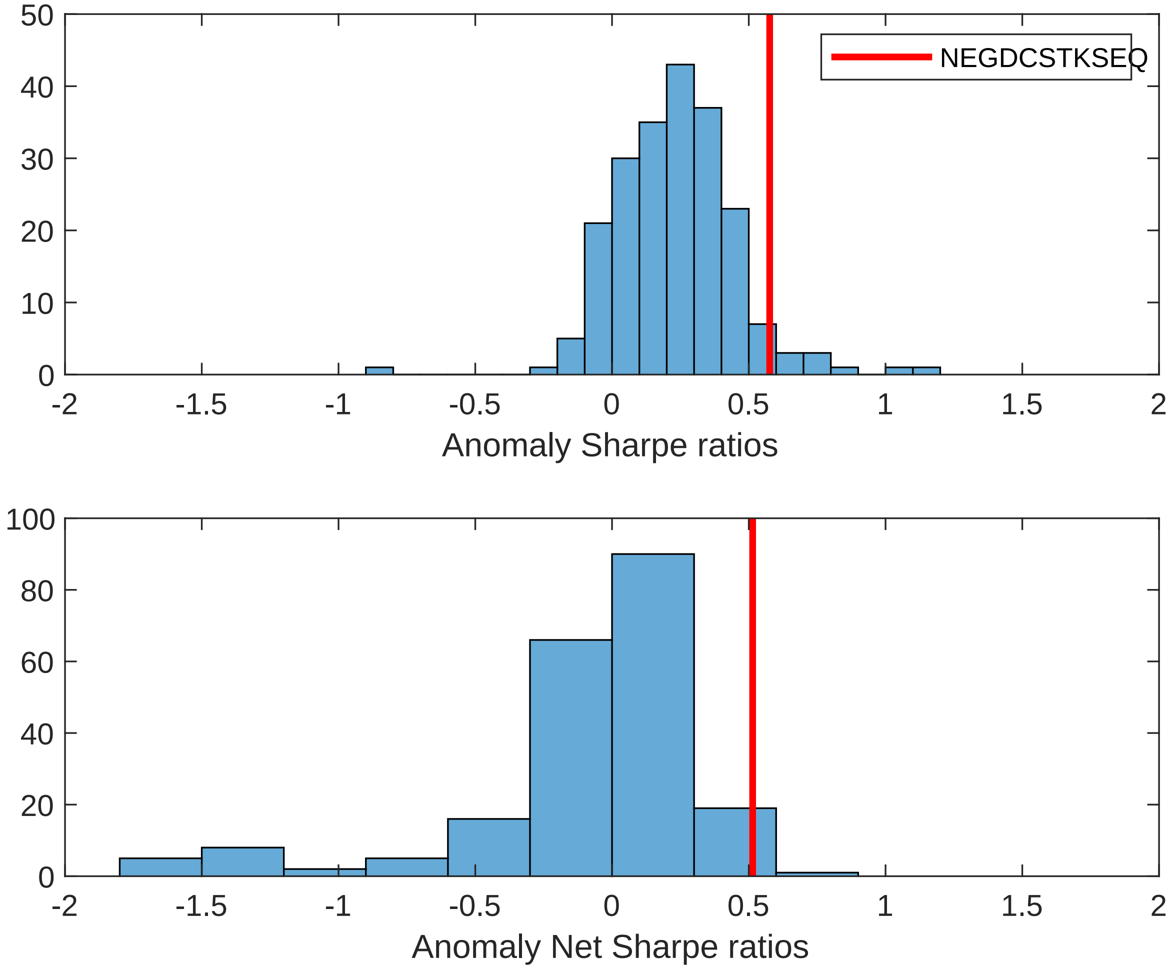

Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the SEIS with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

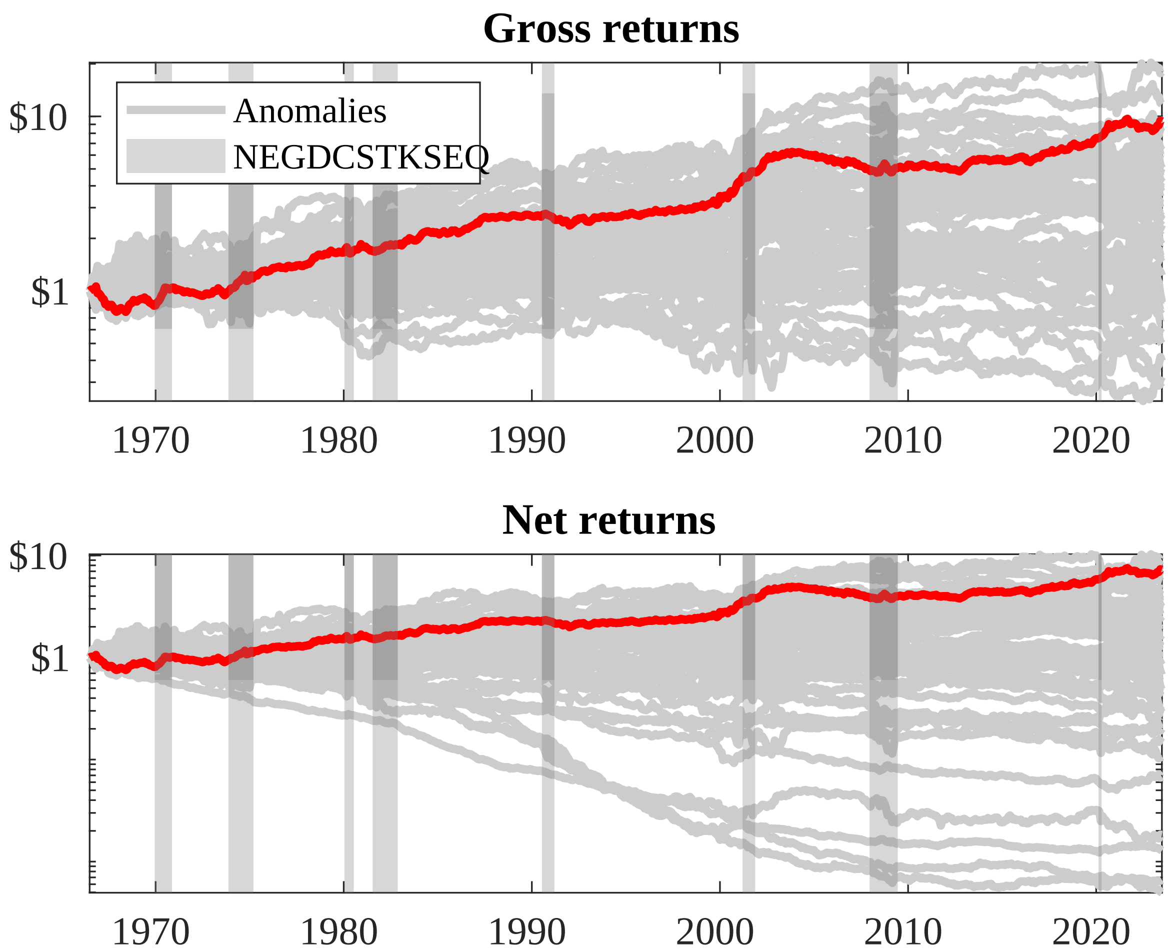

Figure 3: Dollar invested.
This figure plots the growth of a $1 invested in 212 anomaly trading strategies (gray lines), and compares those with the SEIS trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

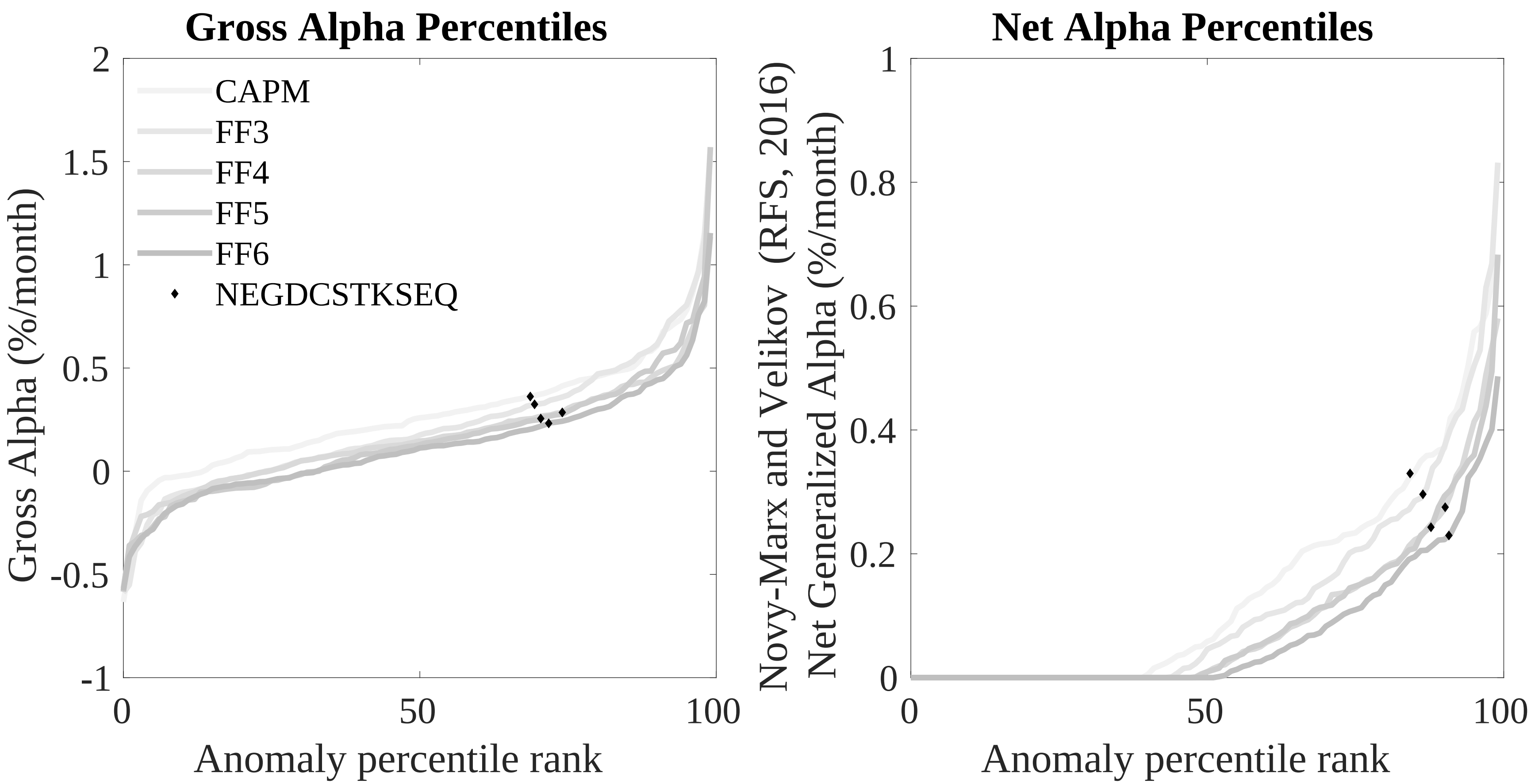

Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the SEIS trading strategy alphas (diamonds). The strategies are constructed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

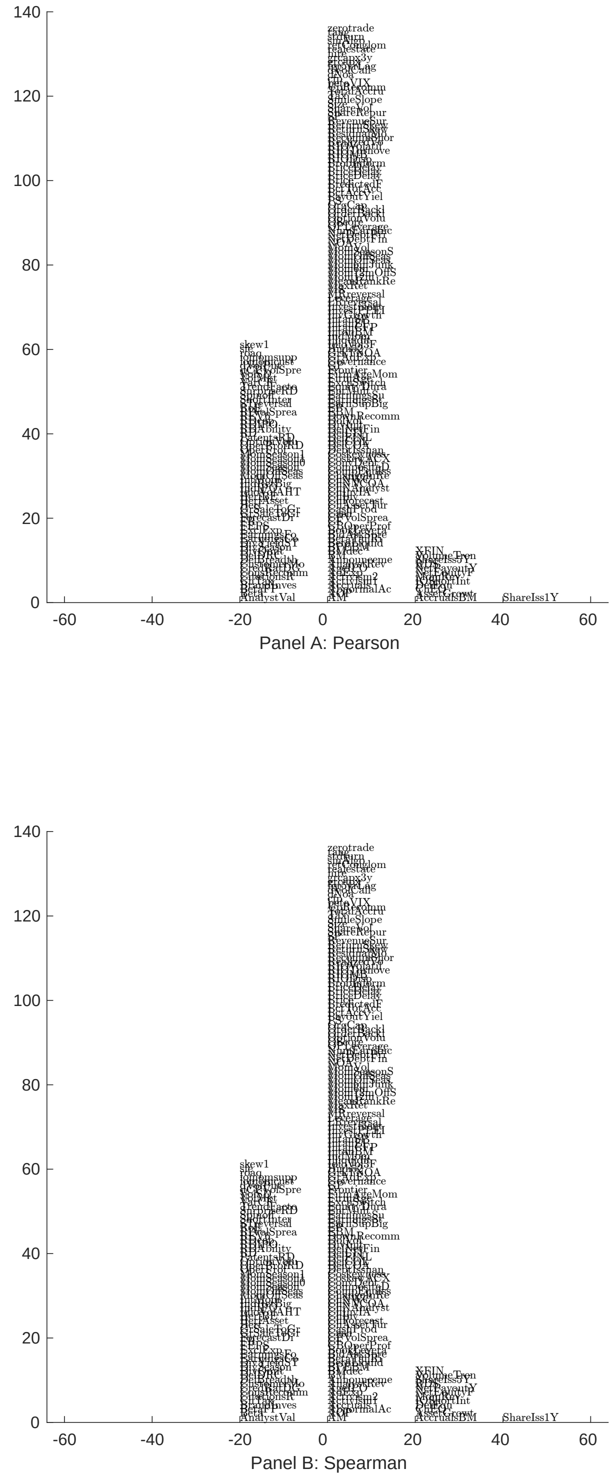

Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 210 filtered anomaly signals with SEIS. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

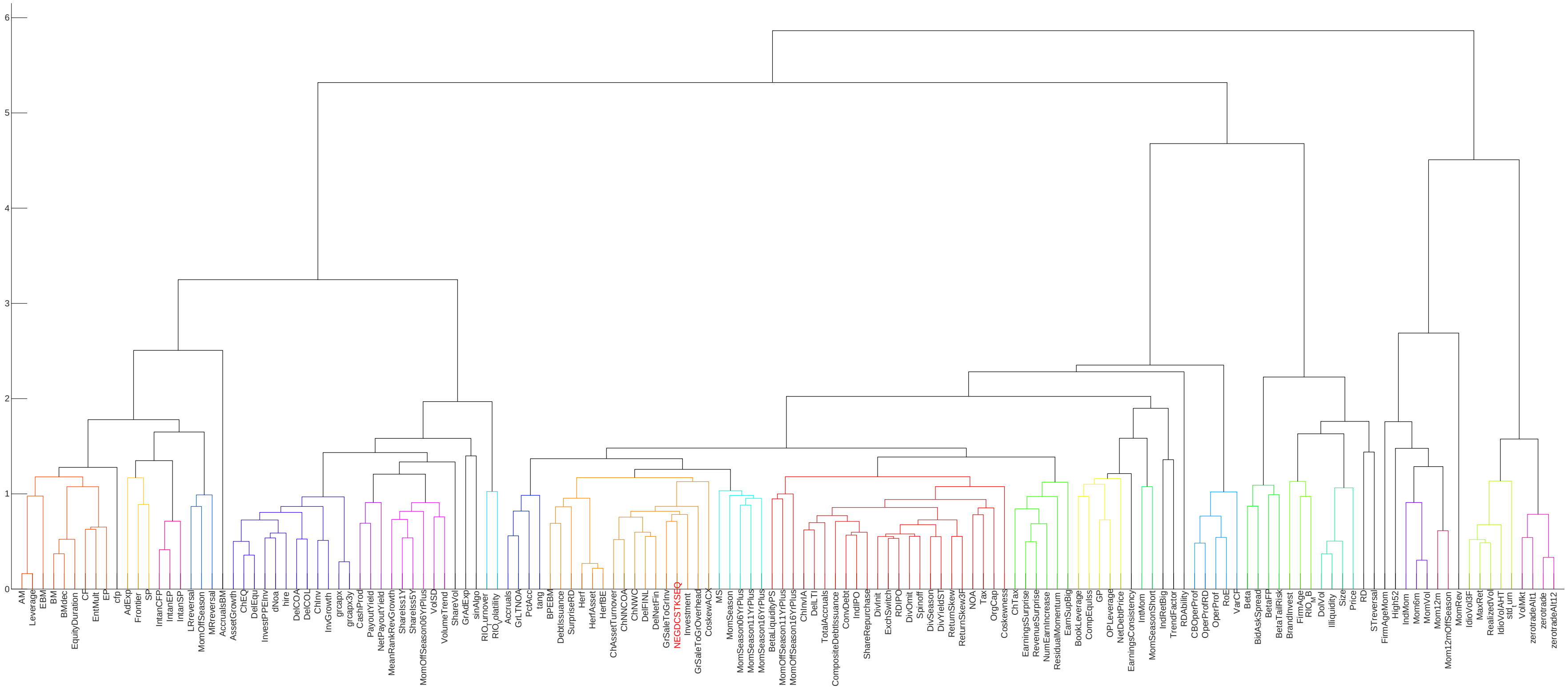

Figure 6: Agglomerative hierarchical cluster plot

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

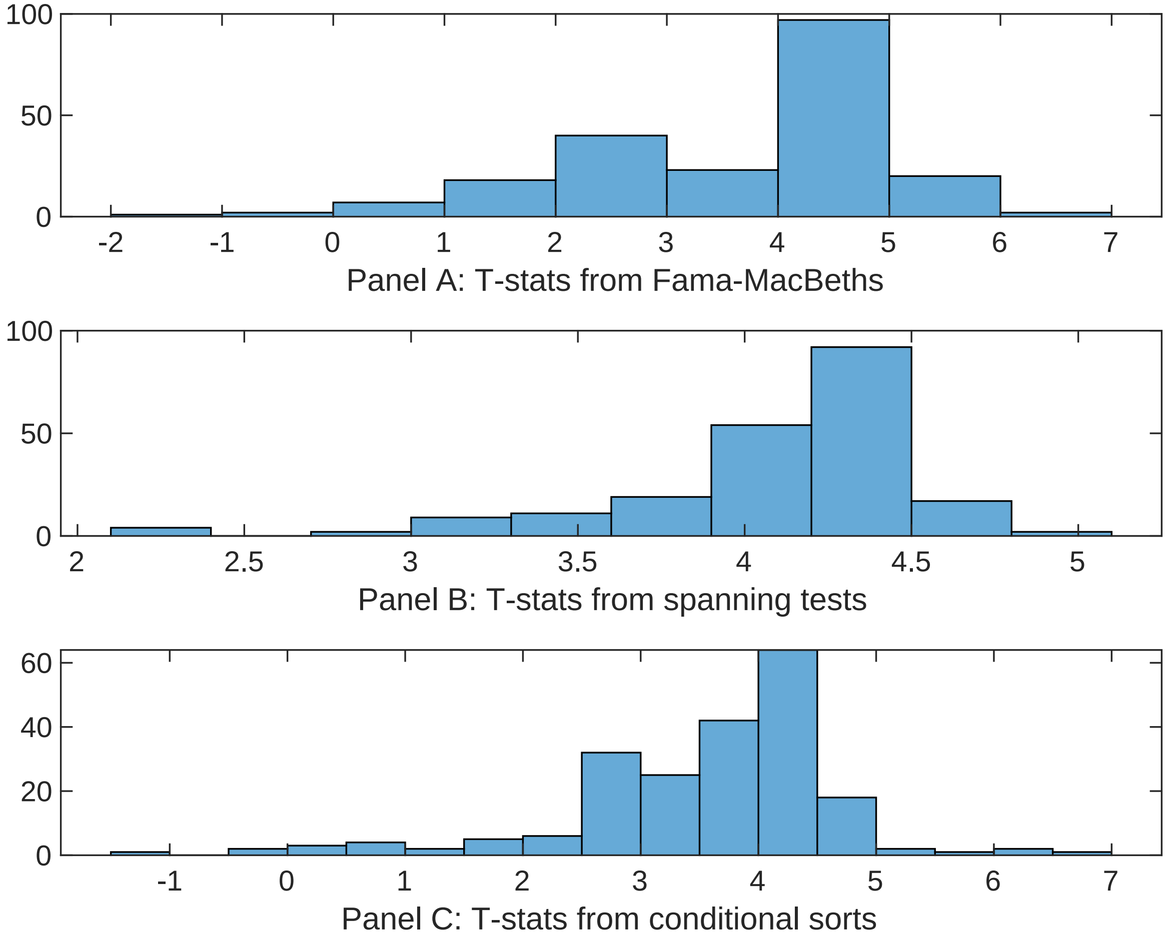

Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of SEIS conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{SEIS} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SEIS} SEIS_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, 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: $r_{SEIS,t} = \alpha + \beta r_{X,t} + \epsilon_t$, 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 SEIS. Stocks are finally grouped into five SEIS portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SEIS 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 SEIS. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{SEIS} SEIS_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. 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	0.17	0.12	0.13	0.12	0.13	0.13
	[5.65]	[7.26]	[5.24]	[6.02]	[5.56]	[6.03]	[5.13]
SEIS	0.23	0.21	0.18	0.24	0.21	0.16	0.13
	[4.31]	[4.24]	[2.88]	[4.34]	[3.92]	[3.09]	[2.11]
Anomaly 1	0.27						0.99
	[5.89]						[2.47]
Anomaly 2		0.47					-0.31
		[4.34]					[-0.21]
Anomaly 3			0.27				0.23
			[2.44]				[2.12]
Anomaly 4				0.38			0.38
				[4.38]			[0.43]
Anomaly 5					0.14		-0.12
					[4.17]		[-0.22]
Anomaly 6						0.10	0.68
						[8.87]	[6.48]
# 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 SEIS trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{SEIS} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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	0.21	0.23	0.22	0.20	0.25	0.24	0.19
	[2.70]	[3.02]	[2.86]	[2.57]	[3.16]	[2.96]	[2.51]
Anomaly 1	27.41						18.92
	[6.94]						[4.16]
Anomaly 2		35.76					40.37
		[8.43]					[6.56]
Anomaly 3			14.55				2.27
			[4.77]				[0.66]
Anomaly 4				14.48			0.92
				[3.51]			[0.21]
Anomaly 5					19.27		-10.70
					[4.61]		[-1.86]
Anomaly 6						3.80	-18.06
						[0.72]	[-3.32]
mkt	3.92	2.89	4.23	3.84	1.41	1.73	5.22
	[2.15]	[1.61]	[2.24]	[2.01]	[0.76]	[0.92]	[2.83]
smb	3.10	0.45	4.80	1.22	1.35	1.31	3.60
	[1.18]	[0.17]	[1.77]	[0.45]	[0.50]	[0.47]	[1.34]
hml	-4.92	-5.94	-6.90	-5.31	-4.18	-1.85	-7.86
	[-1.39]	[-1.70]	[-1.82]	[-1.39]	[-1.15]	[-0.51]	[-2.12]
rmw	-10.10	0.53	-9.24	-3.77	0.57	-1.43	-7.43
	[-2.68]	[0.15]	[-2.31]	[-1.02]	[0.16]	[-0.39]	[-1.80]
cma	17.08	-5.64	19.87	26.16	9.80	25.24	12.66
	[3.07]	[-0.85]	[3.42]	[4.74]	[1.43]	[3.03]	[1.58]
umd	3.49	3.32	5.08	3.96	4.27	3.77	2.34
	[1.95]	[1.86]	[2.77]	[2.16]	[2.31]	[2.00]	[1.32]
# months	680	684	680	680	684	684	680
$\bar{R}^2(\%)$	15	16	12	11	10	7	21

Panel A: Average rank

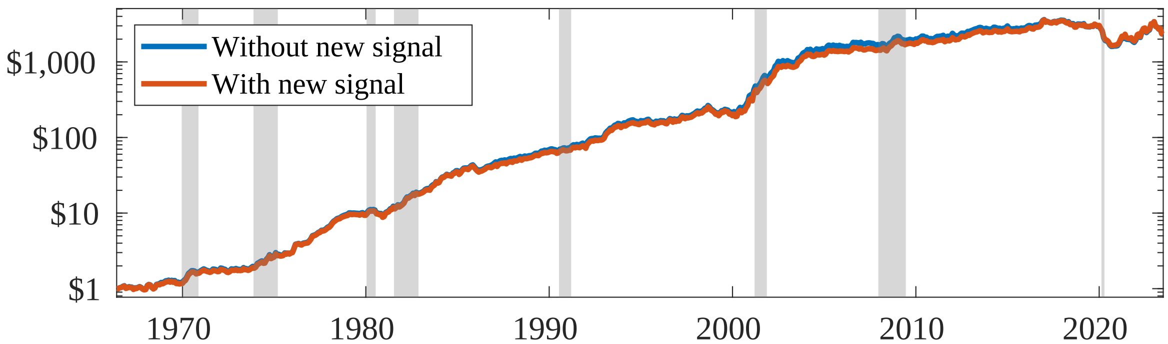

Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 155 anomalies. The red solid lines indicate combination trading strategies that utilize the 155 anomalies as well as SEIS. 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. *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.

Cochrane, J. H. and Pedersen, L. H. (2023). What is the factor zoo? *Review of Financial Studies*, 36(1):1–44.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *Journal of Finance*, 61(4):1605–1643.

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.

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.

Myers, S. C. (1984). The capital structure puzzle. *Journal of Finance*, 39(3):575–592.

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. *Journal of Finance*, 63(2):921–945.