Stock-Intangible Disparity and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Stock-Intangible Disparity (SID), 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 SID achieves an annualized gross (net) Sharpe ratio of 0.52 (0.47), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 21 (21) bps/month with a t-statistic of 2.49 (2.48), 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, Share issuance (5 year), Growth in book equity, Change in equity to assets, Asset growth) is 21 bps/month with a t-statistic of 2.56.

1 Introduction

Market efficiency remains a central question in asset pricing, with mounting evidence that certain firm characteristics can predict future stock returns. While the literature has extensively documented various return predictors, fundamental questions persist about how firms' financing and investment decisions affect their market valuations. Recent work has highlighted the growing importance of intangible capital in modern economies, yet our understanding of how markets price the relationship between firms' equity and their intangible assets remains incomplete. This gap is particularly notable given that intangible capital now comprises over 50% of corporate investment in developed economies.

Prior research has largely focused on either equity issuance or intangible investment in isolation, without fully exploring how their interaction affects asset prices. While Daniel and Titman (2006) show that equity issuance negatively predicts returns and Peters and Taylor (2016) document the pricing implications of intangible investment, the literature has not systematically examined whether the disparity between these two corporate decisions contains incremental information about future stock returns.

We propose that the disparity between a firm's equity issuance and its intangible capital accumulation (Stock-Intangible Disparity or SID) provides a novel signal about future stock returns. Our hypothesis builds on two established theoretical frameworks. First, the market timing theory of Baker and Wurgler (2002) suggests that managers issue equity when they believe their stock is overvalued. Second, the q-theory of investment (Cochrane (1995)) implies that firms optimally invest in intangible capital when they have valuable growth opportunities.

When firms issue equity without corresponding increases in intangible investment, this may indicate opportunistic market timing by management rather than financing of valuable growth opportunities. Conversely, when firms fund intangible investment through means other than equity issuance, this could signal management's confidence in project value and unwillingness to dilute existing shareholders. This framework suggests that firms with high equity issuance relative to intangible investment (high SID) should underperform firms with low equity issuance relative to intangible investment (low SID).

This mechanism is distinct from but complementary to existing explanations for the negative relationship between equity issuance and returns (Pontiff and Woodgate (2008)) and the investment-return relationship (Titman et al. (2004)). By capturing the interaction between these corporate decisions, SID may provide incremental information about managerial views of firm value and investment opportunities.

Our empirical analysis strongly supports the predictive power of Stock-Intangible Disparity. A value-weighted long-short portfolio that buys stocks with low SID and sells stocks with high SID generates a monthly alpha of 21 basis points (t-statistic = 2.49) relative to the Fama-French six-factor model. The strategy achieves an annualized gross (net) Sharpe ratio of 0.52 (0.47), placing it in the top 6% of documented return predictors.

Importantly, the predictive power of SID remains robust across various methodological choices and controls. The signal maintains significant predictability even among large-cap stocks, with a monthly alpha of 20 basis points (t-statistic = 1.88) for stocks above the 80th percentile of market capitalization. This suggests that the effect is not confined to small, illiquid stocks where trading costs might impede implementation.

Most notably, SID's predictive power persists after controlling for the six most closely related anomalies, including share issuance, asset growth, and investment-based factors. In spanning tests that control for these related anomalies and the Fama-French six factors simultaneously, the SID strategy still generates a monthly alpha of 21 basis points (t-statistic = 2.56), indicating that it captures unique information about future returns.

Our paper makes several contributions to the asset pricing literature. First, we introduce a novel predictor that captures the joint information in firms' financing and investment decisions, extending work by Pontiff and Woodgate (2008) on equity issuance and Titman et al. (2004) on corporate investment. While these decisions have been studied separately, we show that their interaction provides incremental information about future returns.

Second, we contribute to the growing literature on intangible capital and asset prices. While Peters and Taylor (2016) examine how intangible investment affects expected returns and Eisfeldt and Papanikolaou (2013) study the risk properties of intangible capital, we demonstrate that the relationship between intangible investment and financing choices helps predict the cross-section of returns. This finding highlights the importance of considering both sources and uses of capital when evaluating firms' investment decisions.

Finally, our results have important implications for both academic research and investment practice. For researchers, we provide new evidence on how markets process information about corporate financing and investment decisions. For practitioners, we document a robust return predictor that remains significant after transaction costs and performs well among large, liquid stocks. The strategy's performance suggests meaningful economic profits from identifying misaligned financing and investment decisions.

2 Data

Our study investigates the predictive power of a financial signal derived from accounting data for cross-sectional returns, focusing specifically on the Stock-Intangible Disparity measure. 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 INTAN for intangible assets. Common stock (CSTK) represents the total value of common shares issued by the company, while intangible assets (INTAN) encompass non-physical assets such as patents, trademarks, goodwill, and other intellectual property. The construction of the signal follows a dynamic difference approach, where we calculate the change in common stock (CSTK minus its lagged value) and scale this difference by the lagged value of intangible assets (INTAN). This scaled difference captures the relative change in equity capital structure in relation to the firm's intellectual capital base. By focusing on this relationship, the signal aims to reflect aspects of capital structure dynamics and intellectual capital utilization in a manner that is both economically meaningful and interpretable. We construct this measure using end-of-fiscal-year values for both CSTK and INTAN to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the SID signal. Panel A plots the time-series of the mean, median, and interquartile range for SID. On average, the cross-sectional mean (median) SID is -0.62 (-0.00) over the 1967 to 2023 sample, where the starting date is determined by the availability of the input SID 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 SID signal for the CRSP universe. On average, the SID signal is available for 3.94% of CRSP names, which on average make up 5.09% of total market capitalization.

4 Does SID predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on SID 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 SID portfolio and sells the low SID 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 SID strategy earns an average return of 0.34% per month with a t-statistic of 3.93. The annualized Sharpe ratio of the strategy is 0.52. The alphas range from 0.21% to 0.38% per month and have t-statistics exceeding 2.49 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 5.10 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 344 stocks and an average market capitalization of at least $1,219 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 31 bps/month with a t-statistics of 3.57. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-five exceed two, and for sixteen 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 28-34bps/month. The lowest return, (28 bps/month), is achieved from the quintile sort using cap breakpoints and value-weighted portfolios, and has an associated t-statistic of 3.17. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the SID 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 SID strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and SID, as well as average returns and alphas for long/short trading SID 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 SID strategy achieves an average return of 29 bps/month with a t-statistic of 2.79. Among these large cap stocks, the alphas for the SID strategy relative to the five most common factor models range from 20 to 33 bps/month with t-statistics between 1.88 and 3.14.

5 How does SID perform relative to the zoo?

Figure 2 puts the performance of SID 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 SID strategy falls in the distribution. The SID strategy's gross (net) Sharpe ratio of 0.52 (0.47) is greater than 94% (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 SID strategy (red line).[2] Ignoring trading costs, a $1 invested in the SID strategy would have yielded $7.32 which ranks the SID strategy in the top 1% across the 212 anomalies. Accounting for trading costs, a $1 invested in the SID strategy would have yielded $5.48 which ranks the SID strategy in the top 1% 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

[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 CAPM, and the Fama-French three-, four-, five-, and six-factor models from Table 1, and indicates the ranking of the SID relative to those. Panel A shows that the SID strategy gross alphas fall between the 65 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 196706 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 SID strategy has a positive net generalized alpha for five out of the five factor models. In these cases SID ranks between the 84 and 90 percentiles in terms of how much it could have expanded the achievable investment frontier.

6 Does SID 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 SID with 208 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 SID or at least to weaken the power SID has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of SID conditioning on each of the 208 filtered

[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.

anomaly signals one at a time. Panel A reports t-statistics on β_{SID} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SID} SID_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 208 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{SID,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 208 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 208 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on SID. Stocks are finally grouped into five SID portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SID trading strategies conditioned on each of the 208 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on SID 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 SID signal in these Fama-MacBeth regressions exceed 0.57, with the minimum t-statistic occurring when controlling for Net Payout Yield. Controlling for all six closely related anomalies, the t-statistic on SID is -0.26.

Similarly, Table 5 reports results from spanning tests that regress returns to the SID 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 SID strategy earns alphas that range from 19-23bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.27, which is achieved when controlling for Net Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the SID trading strategy achieves an alpha of 21bps/month with a t-statistic of 2.56.

7 Does SID add relative to the whole zoo?

Finally, we can ask how much adding SID 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 157 anomalies from the zoo that satisfy our inclusion criteria (blue lines) or these 157 anomalies augmented with the SID 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 157-anomaly combination strategy grows to $2199.16, while $1 investment in the combination strategy that includes SID grows to $3025.95.

8 Conclusion

This study provides compelling evidence for the effectiveness of Stock-Intangible Disparity (SID) as a robust predictor of cross-sectional stock returns. Our findings demonstrate that a value-weighted long/short trading strategy based on SID generates economically and statistically significant returns, with an impressive annualized Sharpe ratio of 0.52 (0.47) on a gross (net) basis. The strategy's persistence in gener-

[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 SID is available.

ating significant abnormal returns, even after controlling for established factors and related anomalies, suggests that SID captures unique information content not fully reflected in current asset pricing models.

Particularly noteworthy is the signal's ability to maintain its predictive power when accounting for transaction costs, as evidenced by the minimal difference between gross and net returns. The robust performance against both the Fama-French five-factor model plus momentum, and an extended model including six closely related anomalies, underscores SID's distinctive contribution to the cross-section of expected returns.

However, several limitations warrant consideration. Our analysis primarily focuses on U.S. equity markets, and the signal's effectiveness in international markets remains to be explored. Additionally, the study period may not fully capture the signal's behavior across different market regimes and economic cycles.

Future research could extend this work in several directions. First, investigating the signal's performance in international markets would provide insights into its global applicability. Second, examining the interaction between SID and other established anomalies could reveal potential complementarities or substitution effects. Finally, exploring the underlying economic mechanisms driving the SID effect would enhance our understanding of this anomaly and its implications for market efficiency.

In conclusion, our findings suggest that Stock-Intangible Disparity represents a valuable addition to the investment practitioner's toolkit, offering meaningful improvements in portfolio performance through its unique information content.

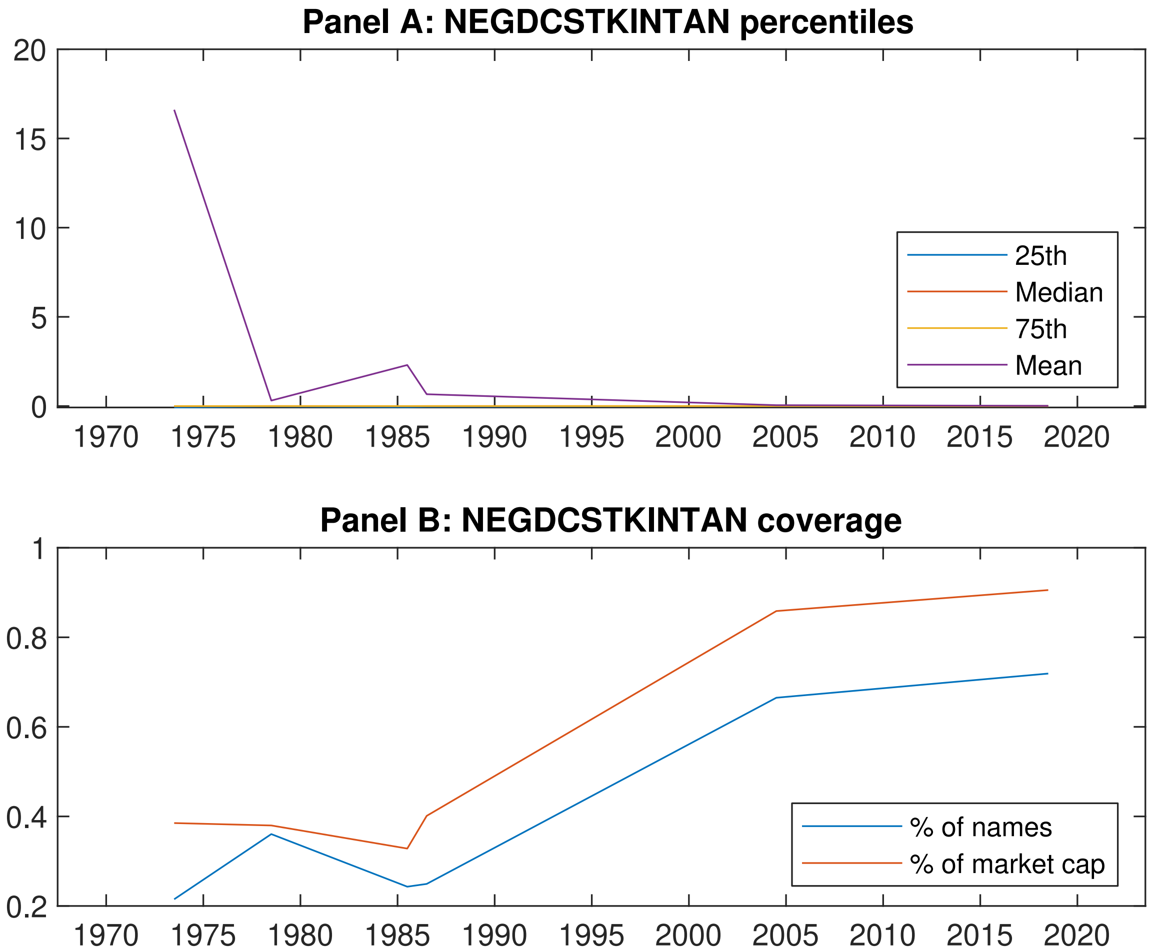

Figure 1: Times series of SID percentiles and coverage.
This figure plots descriptive statistics for SID. Panel A shows cross-sectional percentiles of SID over the sample. Panel B plots the monthly coverage of SID relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on SID. 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 196706 to 202306.

Panel A: Excess returns and alphas on SID-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.45	0.53	0.63	0.72	0.79	0.34
	[2.37]	[2.72]	[3.34]	[4.04]	[4.42]	[3.93]
α_{CAPM}	-0.13	-0.07	0.06	0.19	0.26	0.38
	[-2.09]	[-1.25]	[0.98]	[2.96]	[3.90]	[4.52]
α_{FF3}	-0.13	-0.10	0.04	0.15	0.19	0.31
	[-2.06]	[-1.80]	[0.73]	[2.54]	[2.97]	[3.77]
α_{FF4}	-0.09	-0.07	0.06	0.14	0.20	0.29
	[-1.41]	[-1.33]	[0.96]	[2.19]	[3.18]	[3.44]
α_{FF5}	-0.18	-0.15	-0.04	0.03	0.04	0.22
	[-2.92]	[-2.74]	[-0.67]	[0.55]	[0.68]	[2.61]
α_{FF6}	-0.14	-0.13	-0.02	0.03	0.07	0.21
	[-2.37]	[-2.30]	[-0.37]	[0.45]	[1.10]	[2.49]
Panel B: Fama and French (2018) 6-factor model loadings for SID-sorted portfolios						
β_{MKT}	1.02	1.06	1.03	1.00	0.99	-0.03
	[71.72]	[82.03]	[71.08]	[70.32]	[69.78]	[-1.59]
β_{SMB}	0.01	0.08	0.05	-0.04	0.07	0.06
	[0.37]	[4.00]	[2.16]	[-1.98]	[3.31]	[2.10]
β_{HML}	-0.00	0.04	-0.01	0.03	0.06	0.06
	[-0.11]	[1.52]	[-0.20]	[0.92]	[2.02]	[1.53]
β_{RMW}	0.20	0.15	0.21	0.22	0.26	0.07
	[7.00]	[5.77]	[7.49]	[7.87]	[9.44]	[1.73]
β_{CMA}	-0.04	0.03	0.06	0.17	0.25	0.29
	[-0.97]	[0.81]	[1.54]	[4.20]	[6.15]	[5.10]
β_{UMD}	-0.05	-0.04	-0.03	0.01	-0.04	0.01
	[-3.68]	[-2.82]	[-1.92]	[0.63]	[-2.89]	[0.58]
Panel C: Average number of firms (n) and market capitalization (me)						
n	477	421	344	390	424	
me ($\$10^6$)	1251	1219	1790	1980	2025	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the SID 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 196706 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	0.38	0.31	0.29	0.22	0.21
			[3.93]	[4.52]	[3.77]	[3.44]	[2.61]	[2.49]
Quintile	NYSE	EW	0.50	0.57	0.49	0.43	0.38	0.33
			[7.58]	[8.97]	[8.30]	[7.20]	[6.47]	[5.74]
Quintile	Name	VW	0.35	0.39	0.32	0.30	0.24	0.23
			[4.04]	[4.55]	[3.80]	[3.51]	[2.77]	[2.67]
Quintile	Cap	VW	0.31	0.35	0.30	0.28	0.25	0.24
			[3.57]	[4.04]	[3.45]	[3.21]	[2.80]	[2.72]
Decile	NYSE	VW	0.38	0.41	0.32	0.30	0.25	0.24
			[3.48]	[3.73]	[3.05]	[2.76]	[2.27]	[2.16]
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	0.35	0.29	0.28	0.21	0.21
			[3.50]	[4.10]	[3.47]	[3.32]	[2.49]	[2.48]
Quintile	NYSE	EW	0.31	0.37	0.29	0.26	0.17	0.15
			[4.29]	[5.28]	[4.57]	[4.09]	[2.70]	[2.47]
Quintile	Name	VW	0.31	0.36	0.30	0.29	0.23	0.23
			[3.62]	[4.14]	[3.52]	[3.38]	[2.66]	[2.66]
Quintile	Cap	VW	0.28	0.32	0.27	0.27	0.23	0.23
			[3.17]	[3.66]	[3.17]	[3.06]	[2.70]	[2.68]
Decile	NYSE	VW	0.34	0.36	0.29	0.28	0.22	0.22
			[3.11]	[3.32]	[2.75]	[2.62]	[2.10]	[2.08]

Table 3: Conditional sort on size and SID

This table presents results for conditional double sorts on size and SID. 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 SID. 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 SID and short stocks with low SID .Panel B documents the average number of firms and the average firm size for each portfolio. The sample period is 196706 to 202306.

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

	SID Quintiles					SID Strategies					
Size quintiles	(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
(1)	0.40	0.63	0.82	0.98	0.93	0.53	0.58	0.53	0.46	0.46	0.40
	[1.54]	[2.34]	[3.04]	[3.53]	[3.85]	[6.23]	[6.95]	[6.41]	[5.49]	[5.41]	[4.78]
(2)	0.66	0.65	0.82	0.81	0.91	0.25	0.30	0.23	0.20	0.19	0.17
	[2.74]	[2.59]	[3.26]	[3.31]	[3.99]	[2.51]	[3.08]	[2.37]	[2.05]	[1.92]	[1.73]
(3)	0.44	0.64	0.80	0.83	0.96	0.52	0.53	0.49	0.51	0.44	0.46
	[1.96]	[2.67]	[3.41]	[3.65]	[4.37]	[5.68]	[5.79]	[5.41]	[5.49]	[4.70]	[4.84]
(4)	0.50	0.64	0.71	0.81	0.83	0.32	0.39	0.32	0.33	0.21	0.23
	[2.33]	[2.94]	[3.23]	[3.83]	[4.15]	[3.48]	[4.34]	[3.66]	[3.71]	[2.34]	[2.52]
(5)	0.43	0.51	0.52	0.57	0.73	0.29	0.33	0.26	0.25	0.21	0.20
	[2.31]	[2.69]	[2.78]	[3.26]	[4.17]	[2.79]	[3.14]	[2.50]	[2.34]	[1.92]	[1.88]

Panel B: Portfolio average number of firms and market capitalization

	SID Quintiles					SID Quintiles				
	Average n					Average market capitalization ($\$10^6$)				
Size quintiles	(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
(1)	209	208	207	206	207	21	22	27	21	20
(2)	69	69	69	68	68	43	43	44	43	43
(3)	51	50	50	50	51	77	77	77	79	80
(4)	43	43	43	43	44	168	168	173	175	178
(5)	42	42	42	42	42	1057	1283	1492	1396	1458

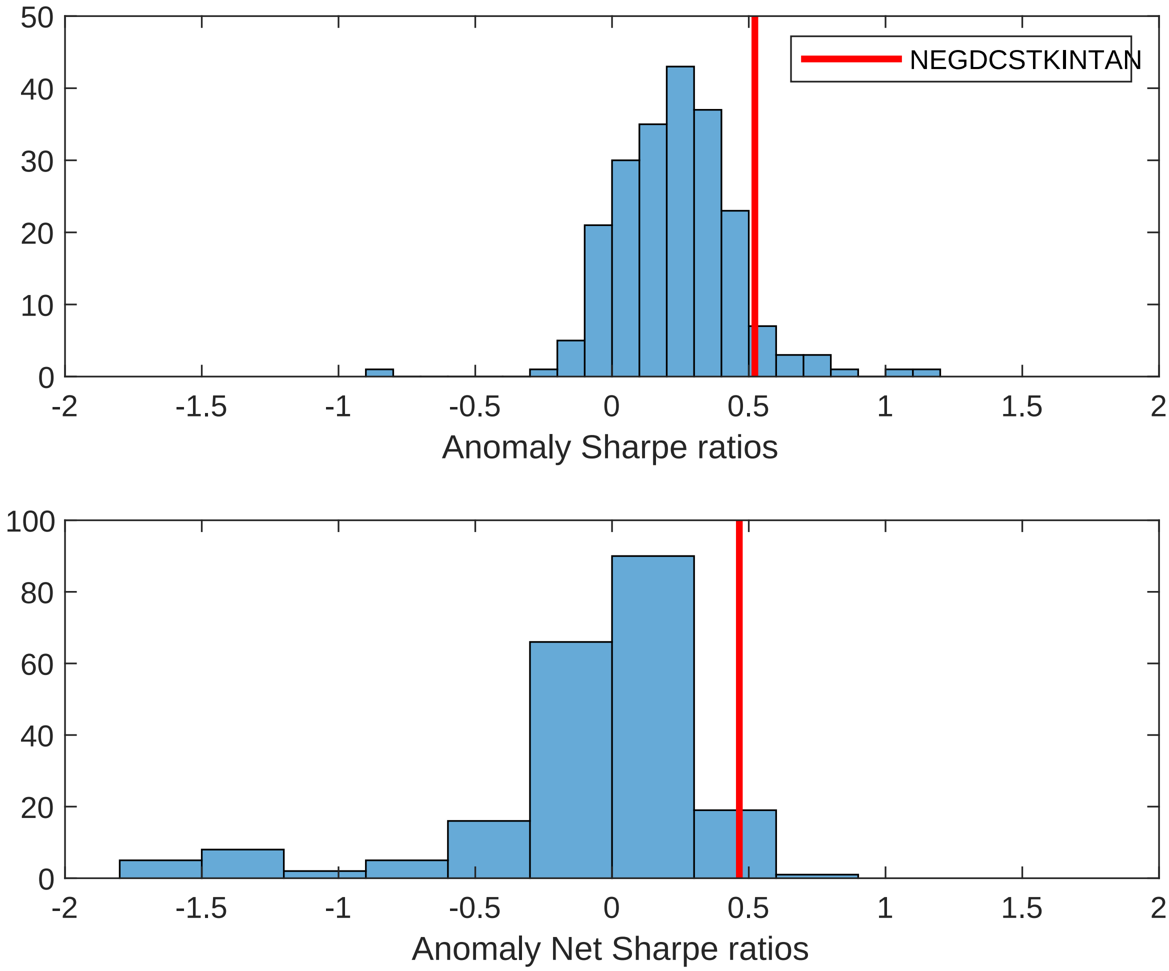

Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the SID with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

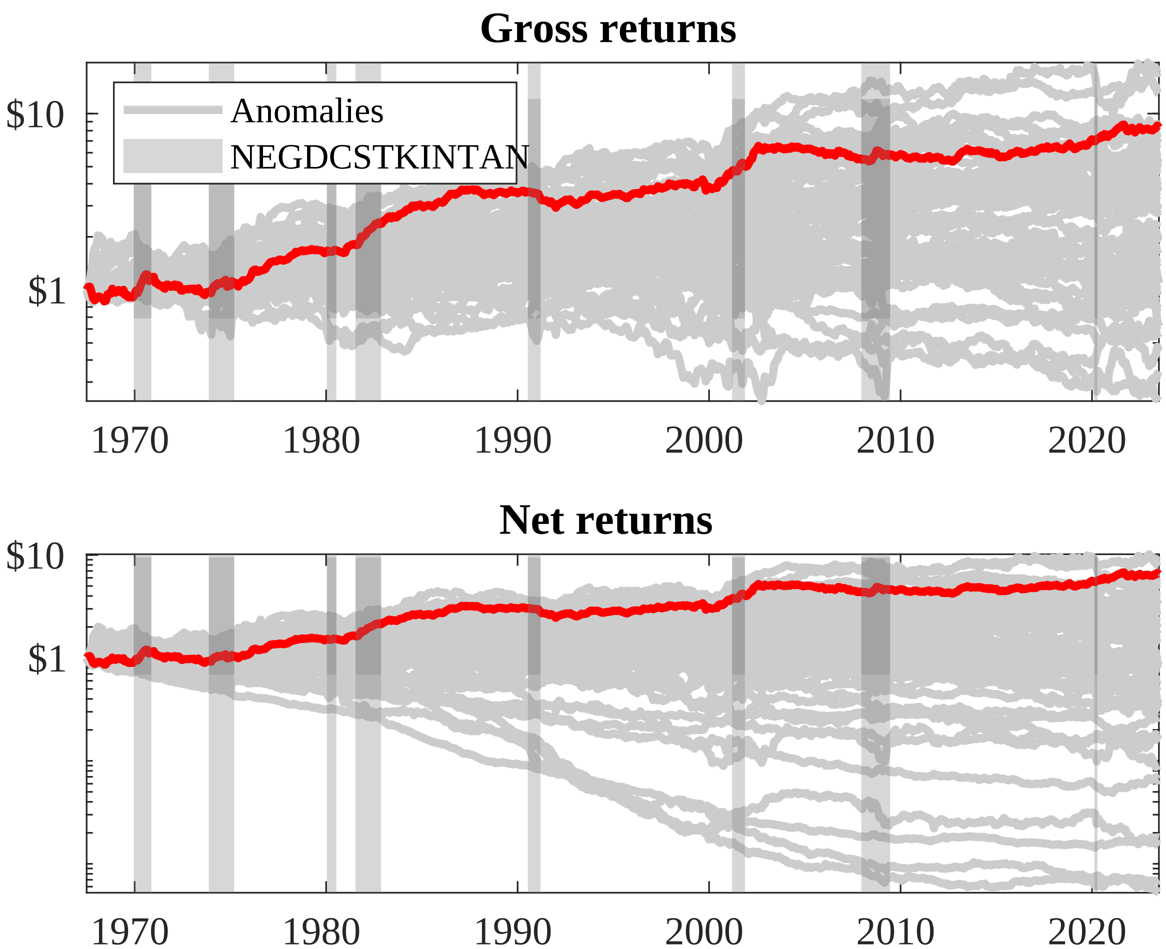

Figure 3: Dollar invested.
This figure plots the growth of a $1 invested in 212 anomaly trading strategies (gray lines), and compares those with the SID trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

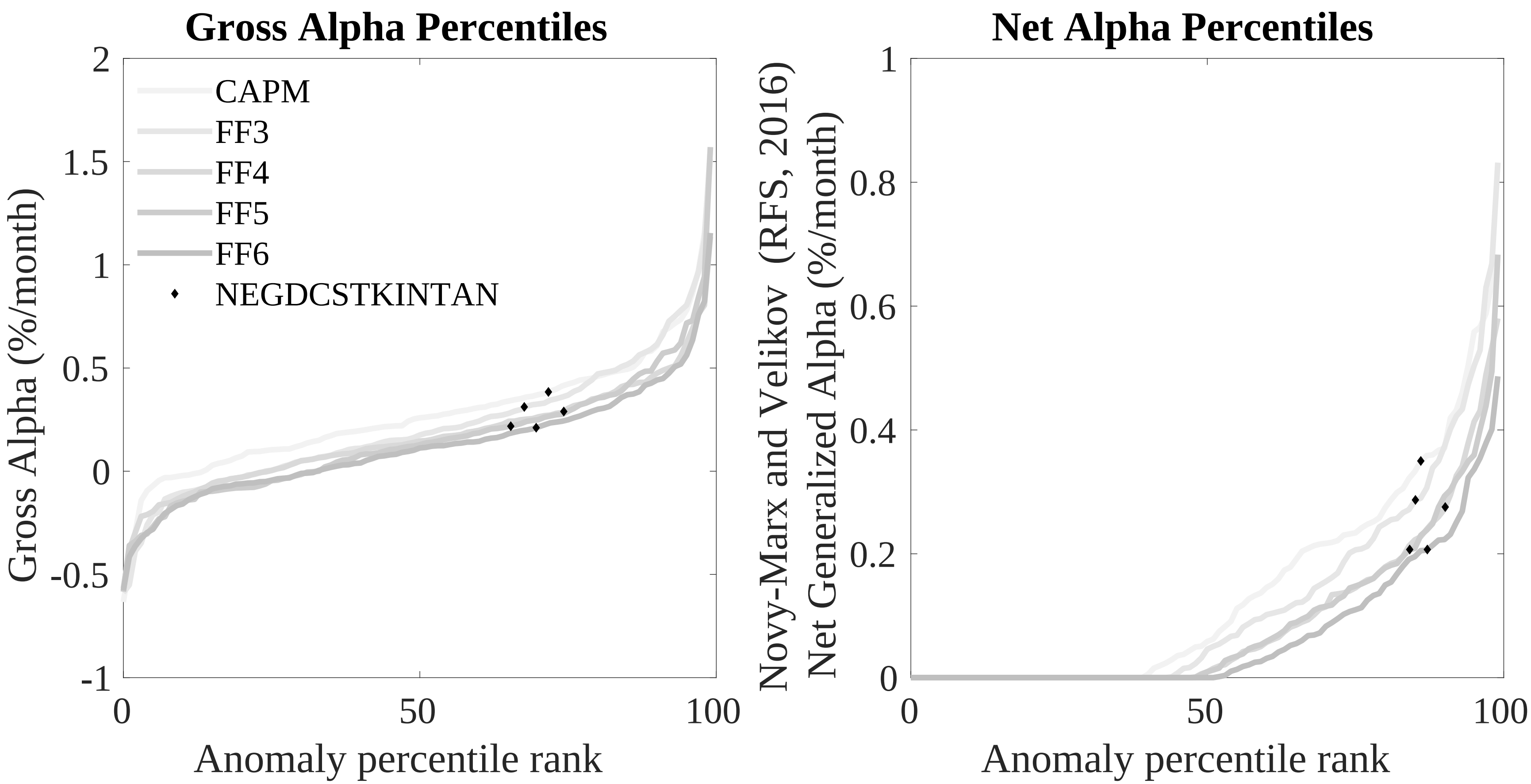

Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the SID trading strategy alphas (diamonds). The strategies are constructed using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

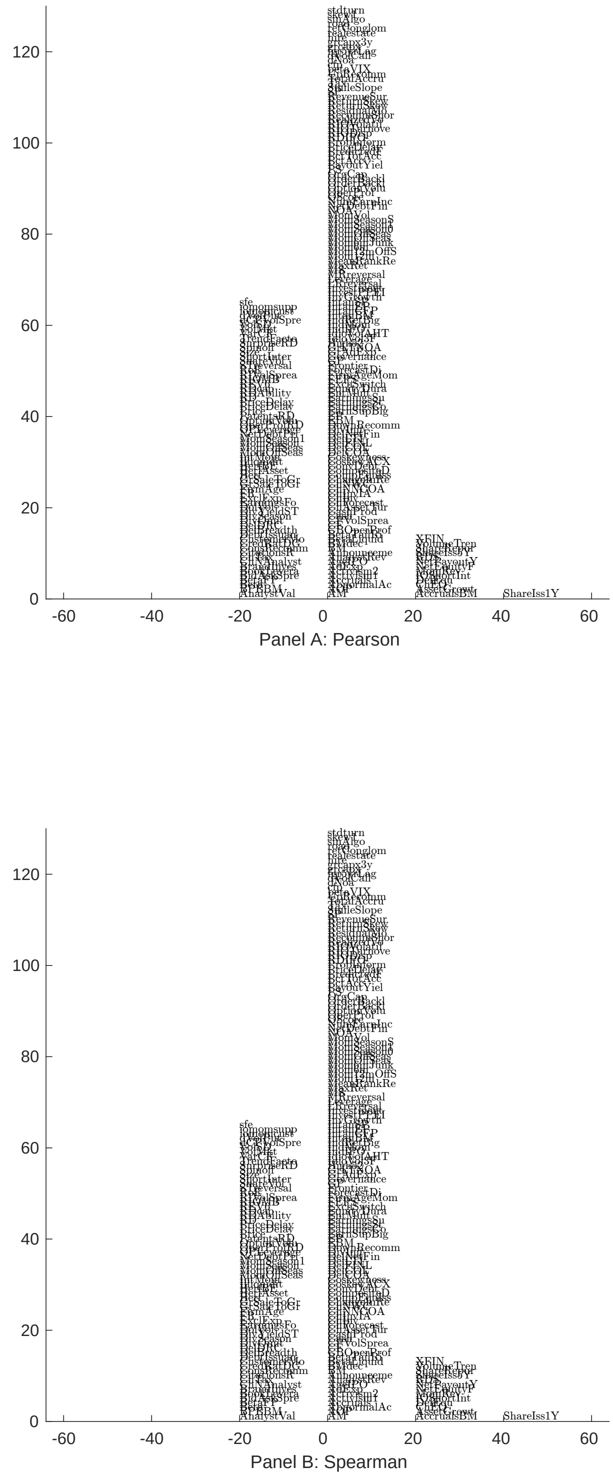

Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 208 filtered anomaly signals with SID. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

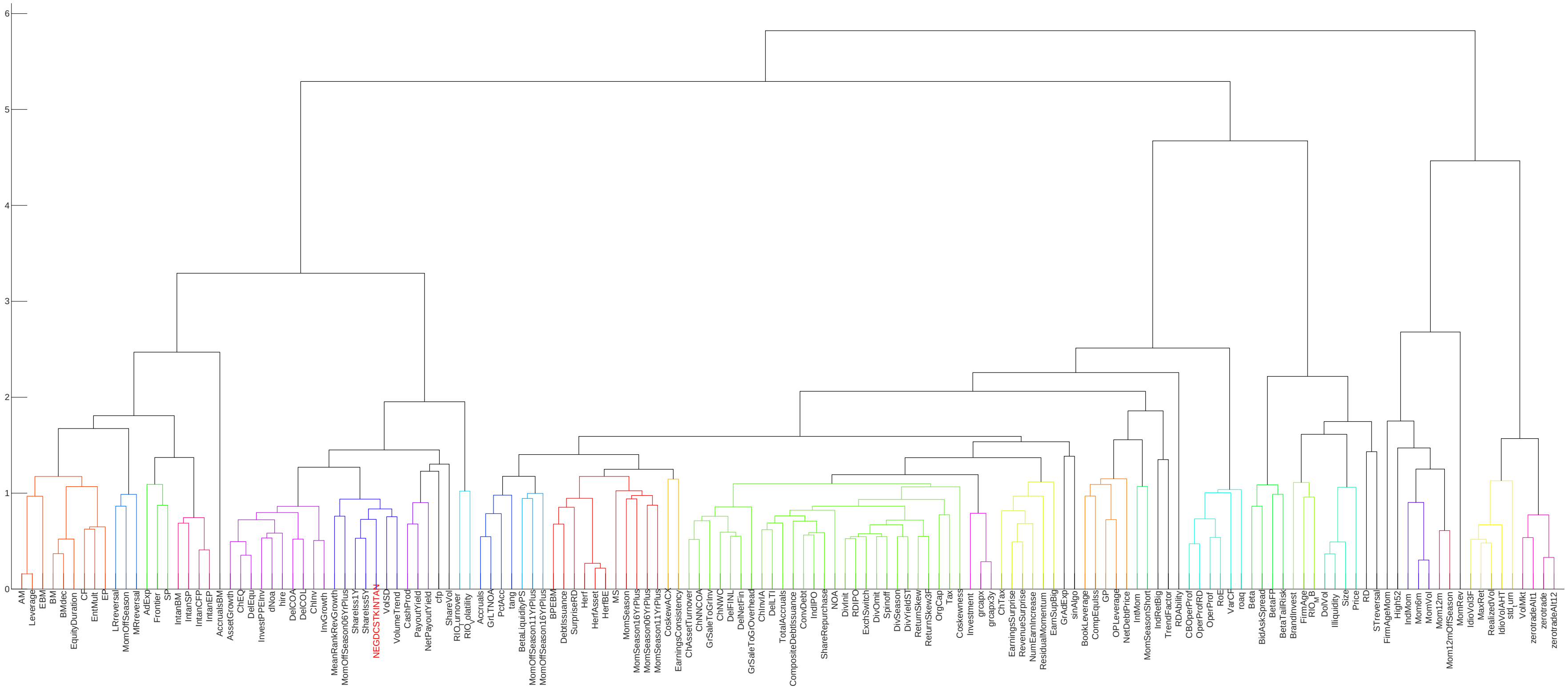

Figure 6: Agglomerative hierarchical cluster plot

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

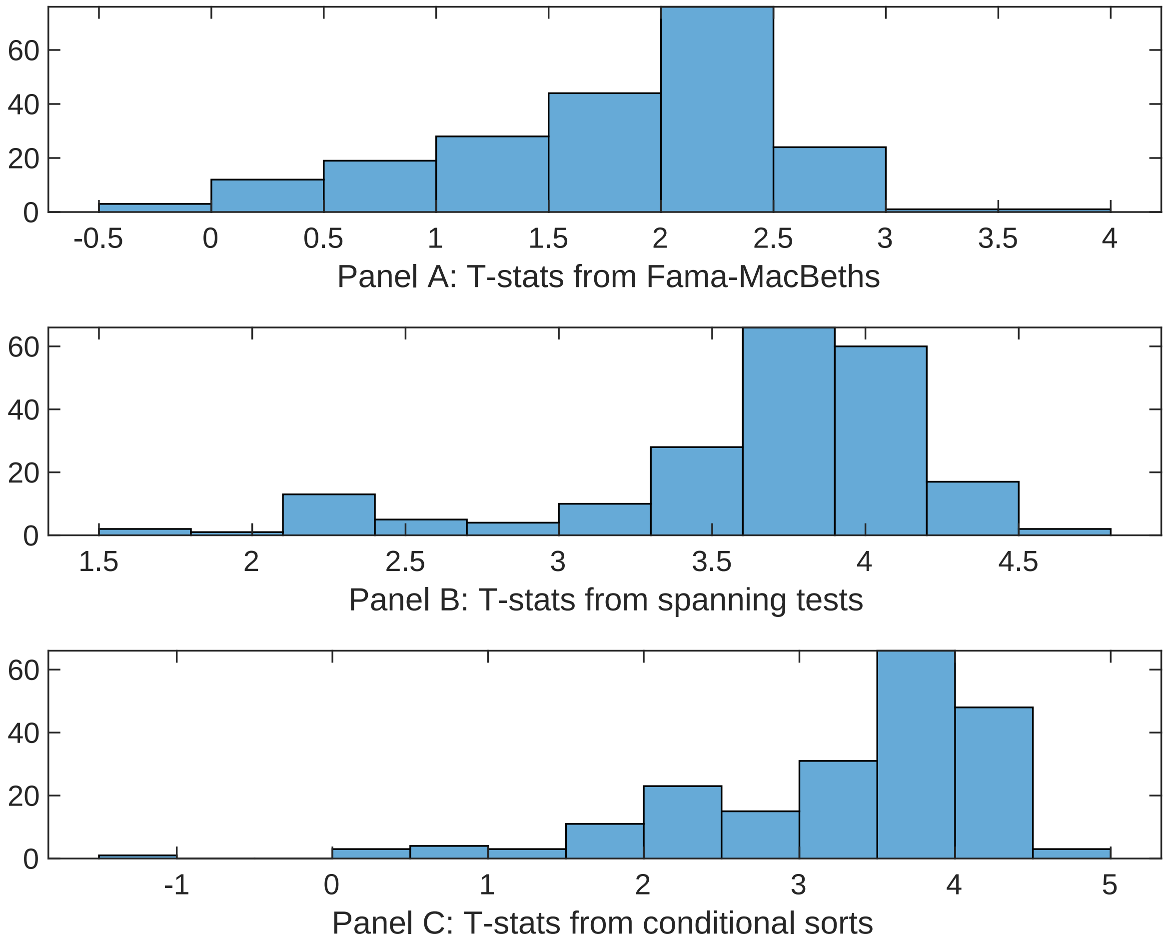

Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of SID conditioning on each of the 208 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{SID} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SID} SID_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 208 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{SID,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 208 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 208 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on SID. Stocks are finally grouped into five SID portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SID trading strategies conditioned on each of the 208 filtered anomalies.

Table 4: Fama-MacBeths controlling for most closely related anomalies
This table presents Fama-MacBeth results of returns on SID. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{SID} SID_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Share issuance (1 year), Net Payout Yield, Share issuance (5 year), Growth in book equity, 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 196706 to 202306.

Intercept	0.12	0.12	0.13	0.18	0.12	0.13	0.12
	[5.34]	[4.95]	[5.70]	[6.83]	[5.31]	[5.76]	[4.73]
SID	0.93	0.35	0.78	0.86	0.86	0.67	-0.16
	[1.79]	[0.57]	[1.45]	[1.92]	[1.70]	[1.36]	[-0.26]
Anomaly 1	0.24						0.54
	[5.03]						[1.13]
Anomaly 2		0.36					0.30
		[3.49]					[2.99]
Anomaly 3			0.35				0.34
			[3.84]				[0.36]
Anomaly 4				0.52			-0.73
				[4.74]			[-0.44]
Anomaly 5					0.15		-0.80
					[4.08]		[-0.12]
Anomaly 6						0.11	0.76
						[8.90]	[6.96]
# months	667	667	667	672	672	672	667
$\bar{R}^2(\%)$	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 SID trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{SID} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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, Share issuance (5 year), Growth in book equity, 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 196706 to 202306.

Intercept	0.19	0.22	0.19	0.21	0.23	0.22	0.21
	[2.40]	[2.71]	[2.27]	[2.51]	[2.73]	[2.63]	[2.56]
Anomaly 1	32.75						18.49
	[7.81]						[3.71]
Anomaly 2		22.62					11.13
		[7.00]					[2.98]
Anomaly 3			24.37				8.96
			[5.63]				[1.91]
Anomaly 4				29.99			15.55
				[6.48]			[2.33]
Anomaly 5					22.48		-2.19
					[5.07]		[-0.35]
Anomaly 6						14.88	-1.17
						[2.68]	[-0.20]
mkt	-0.32	0.84	0.33	-1.88	-3.18	-2.86	1.98
	[-0.17]	[0.42]	[0.16]	[-0.97]	[-1.62]	[-1.44]	[0.99]
smb	7.42	10.64	5.01	5.44	6.03	4.99	7.84
	[2.65]	[3.70]	[1.74]	[1.90]	[2.09]	[1.68]	[2.69]
hml	1.80	-2.48	-0.59	2.86	3.50	5.66	-5.05
	[0.48]	[-0.62]	[-0.15]	[0.76]	[0.92]	[1.47]	[-1.27]
rmw	-4.69	-6.46	1.86	7.86	8.51	6.16	-7.44
	[-1.17]	[-1.53]	[0.48]	[2.07]	[2.20]	[1.58]	[-1.67]
cma	13.50	12.01	20.60	-0.80	5.57	10.89	-3.45
	[2.28]	[1.95]	[3.49]	[-0.11]	[0.76]	[1.23]	[-0.40]
umd	0.65	2.90	1.30	0.76	1.83	1.67	1.33
	[0.34]	[1.51]	[0.67]	[0.39]	[0.94]	[0.84]	[0.70]
# months	668	668	668	672	672	672	668
$\bar{R}^2(\%)$	21	20	18	18	16	14	24

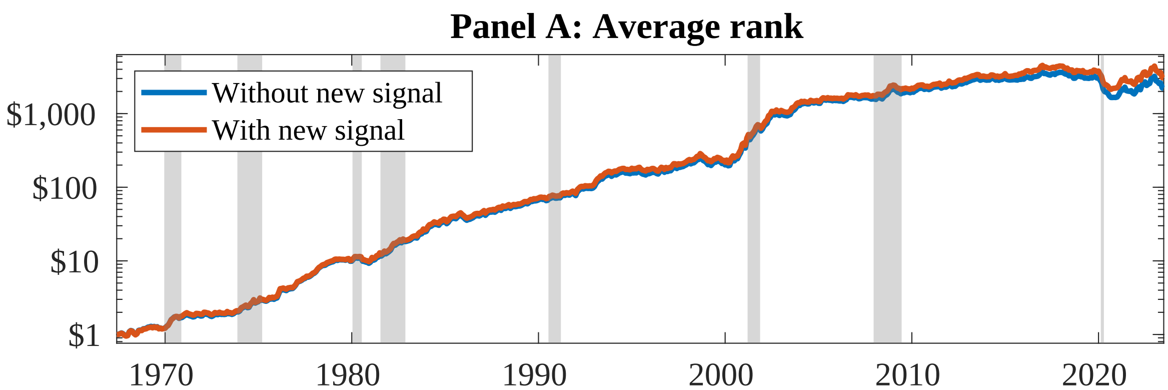

Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 157 anomalies. The red solid lines indicate combination trading strategies that utilize the 157 anomalies as well as SID. 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. (1995). The q-theory of investment. *Journal of Political Economy*, 103(2):242–280.

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*.

Eisfeldt, A. L. and Papanikolaou, D. (2013). Organization capital and the cross-section of expected returns. *Journal of Finance*, 68(4):1365–1406.

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.

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*.

Peters, R. H. and Taylor, L. A. (2016). Intangible capital and the investment-q relation. *Journal of Financial Economics*, 123(2):251–272.

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

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