Tax-Adjusted Stock Difference and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Tax-Adjusted Stock Difference (TSD), 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 TSD achieves an annualized gross (net) Sharpe ratio of 0.46 (0.40), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 18 (16) bps/month with a t-statistic of 2.21 (2.01), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (Net Payout Yield, Share issuance (1 year), Growth in book equity, Share issuance (5 year), Change in equity to assets, Asset growth) is 16 bps/month with a t-statistic of 2.00.

1 Introduction

Market efficiency remains a central question in asset pricing, with substantial debate around whether and how quickly security prices incorporate available information. While classical theory suggests that prices should rapidly reflect all public information, a growing body of evidence documents persistent return predictability from various firm characteristics and signals. One particularly understudied area is how the tax treatment of corporate activities affects stock prices and returns. Despite the economic significance of taxation for firm value, we have limited understanding of how investors process and react to tax-related information.

Prior research has largely focused on broad measures of tax avoidance or specific tax events, leaving open the question of how ongoing differences in firms' tax positions influence their stock returns. This gap is notable given that differences in effective tax rates across firms can materially impact after-tax cash flows and therefore equity values. The relative opacity of corporate tax planning and complexity of tax accounting create potential frictions in how this information is incorporated into prices.

We propose that systematic differences in firms' tax positions, as captured by our Tax-Adjusted Stock Difference (TSD) measure, contain valuable information about future returns. Our hypothesis builds on two key theoretical frameworks. First, following Grossman and Stiglitz (1980), we expect that costly information acquisition leads to only partial incorporation of complex tax information into prices. The technical nature of tax accounting and opacity of tax planning activities suggest that fully processing tax-related information requires significant expertise and resources.

Second, drawing on DellaVigna and Pollet (2009), we argue that limited investor attention particularly affects the processing of tax information given its complexity and indirect relation to fundamental value. When attention is scarce, investors may focus on more salient firm characteristics while underweighting tax-related information. This creates systematic mispricing that only gradually corrects as the implications for cash flows become more apparent.

Finally, building on Hanlon and Heitzman (2010), we hypothesize that firms' tax positions contain forward-looking information about their future profitability and risk. Differences in effective tax rates often reflect underlying economic choices about investment, financing, and operational decisions that have persistent effects on firm value. Therefore, our TSD measure may capture both mispricing of current tax information and predictive signals about future firm performance.

Our empirical analysis reveals strong evidence that TSD predicts future stock returns. A value-weighted long-short portfolio strategy based on TSD quintiles generates significant abnormal returns of 18 basis points per month (t-statistic = 2.21) relative to the Fama-French six-factor model. The strategy achieves an annualized gross Sharpe ratio of 0.46, placing it in the top 11% of documented return predictors.

Importantly, the predictive power of TSD remains robust after controlling for transaction costs and various methodological choices. The strategy continues to earn significant net returns of 16 basis points monthly (t-statistic = 2.01) after accounting for trading frictions. This performance persists across different portfolio construction approaches and remains economically meaningful among large-cap stocks, with the largest size quintile generating abnormal returns of 26 basis points monthly (t-statistic = 2.62).

Further analysis demonstrates that TSD's predictive power is distinct from known anomalies. Controlling for the six most closely related predictors and standard risk factors simultaneously, TSD continues to generate significant abnormal returns of 16 basis points monthly (t-statistic = 2.00). This indicates that TSD captures unique information not reflected in existing return predictors.

Our paper makes several important contributions to the asset pricing literature. First, we extend the work of Hanlon and Heitzman (2010) and Graham and Tucker

(2006) by showing how differences in firms' tax positions systematically predict future returns. While prior research has focused primarily on accounting and real effects of taxation, we demonstrate its importance for asset prices and portfolio choice.

Second, we contribute to the growing literature on return prediction and market efficiency. Our findings complement studies like Hou et al. (2020) by identifying a novel source of predictable returns based on tax information. The persistence and robustness of the TSD effect suggests meaningful frictions in how markets process complex accounting information.

Finally, our results have important implications for both academic research and investment practice. For researchers, we provide new evidence on the links between corporate taxation and asset prices, suggesting fertile ground for future work. For practitioners, our findings identify a new source of potential alpha that is robust to transaction costs and implementable even among large-cap stocks. The relatively low correlation of TSD with existing factors makes it particularly valuable for portfolio diversification.

2 Data

Our study investigates the predictive power of a financial signal derived from accounting data for cross-sectional returns, focusing specifically on the Tax-Adjusted Stock Difference. 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 TXDITC for deferred taxes and investment tax credit. Common stock (CSTK) represents the total value of common shares outstanding, while TXDITC captures the cumulative tax deferrals and credits that arise from timing differences between financial and tax accounting.construction of the signal follows a difference-based approach, where we first calculate the year-over-year change in CSTK (CSTK minus its lagged value) and then scale this difference by the lagged value of TXDITC for each firm in each year of our sample. This scaled difference captures the relative change in common stock value adjusted for the firm's tax position, potentially offering insight into how changes in equity structure interact with tax considerations. By focusing on this relationship, the signal aims to reflect aspects of capital structure decisions and tax planning in a manner that is both scalable and interpretable. We construct this measure using end-of-fiscal-year values for both CSTK and TXDITC to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the TSD signal. Panel A plots the time-series of the mean, median, and interquartile range for TSD. On average, the cross-sectional mean (median) TSD is -0.75 (-0.00) over the 1966 to 2023 sample, where the starting date is determined by the availability of the input TSD data. The signal's interquartile range spans -0.35 to 0.00. Panel B of Figure 1 plots the time-series of the coverage of the TSD signal for the CRSP universe. On average, the TSD signal is available for 3.48% of CRSP names, which on average make up 6.01% of total market capitalization.

4 Does TSD predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on TSD 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 TSD portfolio and sells the low TSD 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 TSD strategy earns an average return of 0.30% per month with a t-statistic of 3.48. The annualized Sharpe ratio of the strategy is 0.46. The alphas range from 0.16% to 0.34% per month and have t-statistics exceeding 1.97 everywhere. The lowest alpha is with respect to the FF5 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 0.28, with a t-statistic of 5.01 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 315 stocks and an average market capitalization of at least $1,088 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 23 bps/month with a t-statistics of 2.74. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-two exceed two, and for thirteen 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 20-29bps/month. The lowest return, (20 bps/month), is achieved from the quintile sort using cap breakpoints and value-weighted portfolios, and has an associated t-statistic of 2.35. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the TSD 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 cases.

Table 3 provides direct tests for the role size plays in the TSD strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and TSD, as well as average returns and alphas for long/short trading TSD 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 TSD strategy achieves an average return of 26 bps/month with a t-statistic of 2.62. Among these large cap stocks, the alphas for the TSD strategy relative to the five most common factor models range from 17 to 29 bps/month with t-statistics between 1.70 and 2.89.

5 How does TSD perform relative to the zoo?

Figure 2 puts the performance of TSD 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 TSD strategy falls in the distribution. The TSD strategy's gross (net) Sharpe ratio of 0.46 (0.40) is greater than 89% (97%) 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 TSD strategy (red line).[2] Ignoring trading costs, a \$1 invested in the TSD strategy would have yielded \$5.55 which ranks the TSD strategy in the top 3% across the 212 anomalies. Accounting for trading costs, a \$1 invested in the TSD strategy would have yielded \$4.12 which ranks the TSD strategy in the top 2% 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 TSD relative to those. Panel A shows that the TSD strategy gross alphas fall between the 56 and 72 percentiles across the five factor models. Panel B shows that, accounting for trading costs, a large fraction of anomalies have not improved the investment opportunity set of an investor with access to the factor models over the 196606 to 202306 sample. For example, 45%

[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.

(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 TSD strategy has a positive net generalized alpha for five out of the five factor models. In these cases TSD ranks between the 75 and 87 percentiles in terms of how much it could have expanded the achievable investment frontier.

6 Does TSD 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 TSD with 207 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 TSD or at least to weaken the power TSD has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of TSD conditioning on each of the 207 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{TSD} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{TSD} TSD_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 207 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{TSD,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 207 filtered anomaly trading strategies at a time. The strategies employed in the spanning tests are constructed using quintile sorts, value-

[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.

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 207 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on TSD. Stocks are finally grouped into five TSD portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted TSD trading strategies conditioned on each of the 207 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on TSD 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 TSD signal in these Fama-MacBeth regressions exceed 2.09, with the minimum t-statistic occurring when controlling for Net Payout Yield. Controlling for all six closely related anomalies, the t-statistic on TSD is 1.14.

Similarly, Table 5 reports results from spanning tests that regress returns to the TSD 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 TSD strategy earns alphas that range from 15-20bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 1.82, which is achieved when controlling for Net Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the TSD trading strategy achieves an alpha of 16bps/month with a t-statistic of 2.00.

7 Does TSD add relative to the whole zoo?

Finally, we can ask how much adding TSD 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 TSD 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 TSD grows to $2143.45.

[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 TSD is available.

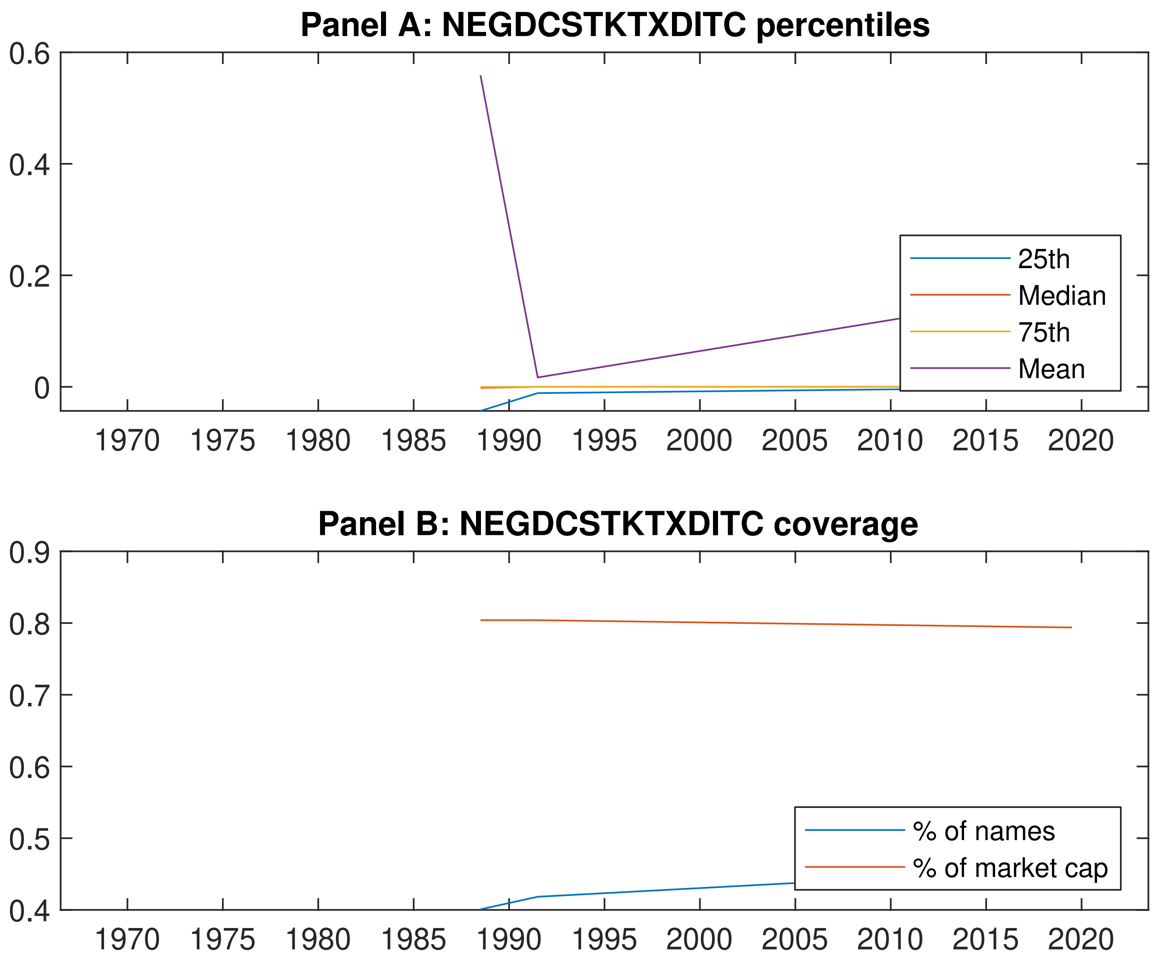


Figure 1: Times series of TSD percentiles and coverage.
This figure plots descriptive statistics for TSD. Panel A shows cross-sectional percentiles of TSD over the sample. Panel B plots the monthly coverage of TSD relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on TSD. 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 TSD-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.45	0.49	0.62	0.67	0.75	0.30
	[2.58]	[2.72]	[3.55]	[4.05]	[4.51]	[3.48]
α_{CAPM}	-0.09	-0.07	0.08	0.16	0.25	0.34
	[-1.67]	[-1.27]	[1.45]	[2.91]	[4.03]	[3.98]
α_{FF3}	-0.06	-0.07	0.06	0.14	0.20	0.26
	[-1.16]	[-1.29]	[1.07]	[2.73]	[3.41]	[3.19]
α_{FF4}	-0.07	-0.03	0.05	0.10	0.20	0.27
	[-1.24]	[-0.61]	[0.92]	[1.97]	[3.48]	[3.30]
α_{FF5}	-0.13	-0.10	-0.05	0.01	0.03	0.16
	[-2.37]	[-1.71]	[-0.88]	[0.16]	[0.60]	[1.97]
α_{FF6}	-0.13	-0.06	-0.04	-0.01	0.05	0.18
	[-2.35]	[-1.14]	[-0.81]	[-0.22]	[0.99]	[2.21]
Panel B: Fama and French (2018) 6-factor model loadings for TSD-sorted portfolios						
β_{MKT}	0.97	0.98	0.99	0.98	0.96	-0.00
	[74.58]	[75.00]	[76.71]	[86.45]	[75.89]	[-0.15]
β_{SMB}	0.01	-0.04	0.01	-0.14	-0.02	-0.03
	[0.53]	[-2.18]	[0.43]	[-8.77]	[-0.98]	[-0.99]
β_{HML}	-0.08	0.00	-0.00	0.00	0.01	0.09
	[-3.35]	[0.15]	[-0.01]	[0.10]	[0.38]	[2.48]
β_{RMW}	0.17	0.09	0.21	0.16	0.28	0.11
	[6.65]	[3.66]	[8.30]	[7.04]	[11.24]	[2.90]
β_{CMA}	0.02	-0.01	0.14	0.29	0.30	0.28
	[0.56]	[-0.24]	[3.84]	[8.90]	[8.26]	[5.01]
β_{UMD}	0.00	-0.05	-0.00	0.03	-0.03	-0.03
	[0.06]	[-3.78]	[-0.37]	[2.57]	[-2.61]	[-1.74]
Panel C: Average number of firms (n) and market capitalization (me)						
n	401	360	315	373	424	
me ($\$10^6$)	1210	1088	1489	1696	1784	

Table 2: Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the TSD 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	0.30	0.34	0.26	0.27	0.16	0.18
			[3.48]	[3.98]	[3.19]	[3.30]	[1.97]	[2.21]
Quintile	NYSE	EW	0.45	0.52	0.43	0.38	0.31	0.29
			[7.01]	[8.65]	[8.12]	[7.20]	[6.11]	[5.54]
Quintile	Name	VW	0.33	0.37	0.29	0.29	0.19	0.20
			[3.87]	[4.38]	[3.60]	[3.50]	[2.36]	[2.42]
Quintile	Cap	VW	0.23	0.27	0.22	0.23	0.17	0.19
			[2.74]	[3.17]	[2.62]	[2.74]	[2.01]	[2.21]
Decile	NYSE	VW	0.30	0.35	0.25	0.24	0.19	0.19
			[2.92]	[3.43]	[2.54]	[2.38]	[1.91]	[1.86]
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	0.26	0.30	0.23	0.24	0.15	0.16
			[3.06]	[3.56]	[2.88]	[2.96]	[1.80]	[2.01]
Quintile	NYSE	EW	0.29	0.36	0.28	0.25	0.15	0.15
			[4.27]	[5.54]	[4.85]	[4.48]	[2.75]	[2.67]
Quintile	Name	VW	0.29	0.34	0.27	0.26	0.18	0.19
			[3.44]	[3.97]	[3.30]	[3.26]	[2.21]	[2.33]
Quintile	Cap	VW	0.20	0.24	0.19	0.20	0.15	0.17
			[2.35]	[2.78]	[2.28]	[2.37]	[1.81]	[1.95]
Decile	NYSE	VW	0.26	0.30	0.21	0.21	0.15	0.16
			[2.52]	[2.95]	[2.17]	[2.10]	[1.54]	[1.61]

Table 3: Conditional sort on size and TSD

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

		TSD Quintiles					TSD Strategies					
		(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Size quintiles	(1)	0.71	0.80	1.00	0.95	1.04	0.33	0.37	0.28	0.26	0.20	0.19
		[2.82]	[3.15]	[3.83]	[3.39]	[4.36]	[3.13]	[3.59]	[2.84]	[2.62]	[1.98]	[1.87]
	(2)	0.59	0.74	0.78	0.95	0.98	0.40	0.48	0.36	0.33	0.29	0.28
		[2.45]	[3.14]	[3.33]	[4.12]	[4.47]	[3.98]	[4.94]	[4.03]	[3.69]	[3.19]	[3.00]
	(3)	0.63	0.68	0.68	0.85	0.99	0.36	0.40	0.32	0.32	0.27	0.27
		[2.93]	[3.21]	[3.07]	[4.06]	[4.88]	[4.24]	[4.80]	[4.00]	[3.89]	[3.25]	[3.25]
	(4)	0.53	0.61	0.76	0.75	0.80	0.26	0.33	0.23	0.18	0.10	0.07
		[2.59]	[3.03]	[3.74]	[3.77]	[4.22]	[3.05]	[3.85]	[2.95]	[2.29]	[1.22]	[0.86]
	(5)	0.45	0.51	0.46	0.56	0.71	0.26	0.29	0.23	0.26	0.17	0.20
		[2.59]	[2.92]	[2.63]	[3.37]	[4.27]	[2.62]	[2.89]	[2.34]	[2.63]	[1.70]	[2.03]

Panel B: Portfolio average number of firms and market capitalization

		TSD Quintiles					TSD Quintiles				
		Average n					Average market capitalization ($\$10^6$)				
		(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
Size quintiles	(1)	165	164	164	163	164	13	14	16	12	12
	(2)	64	64	63	64	63	31	31	30	31	31
	(3)	52	52	52	52	52	61	61	61	63	63
	(4)	49	49	49	49	49	145	144	151	151	150
	(5)	46	46	46	46	46	1033	1160	1302	1198	1305

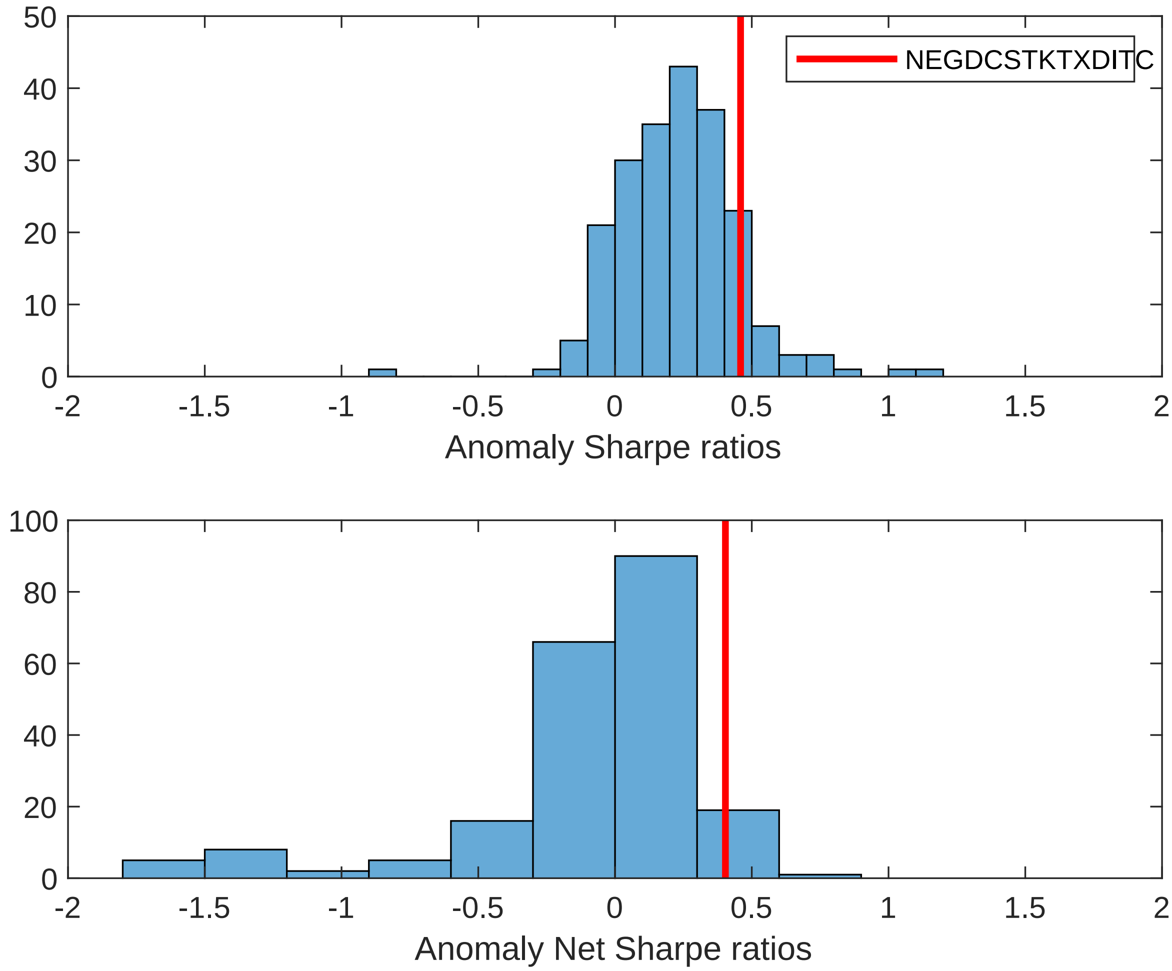


Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the TSD with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

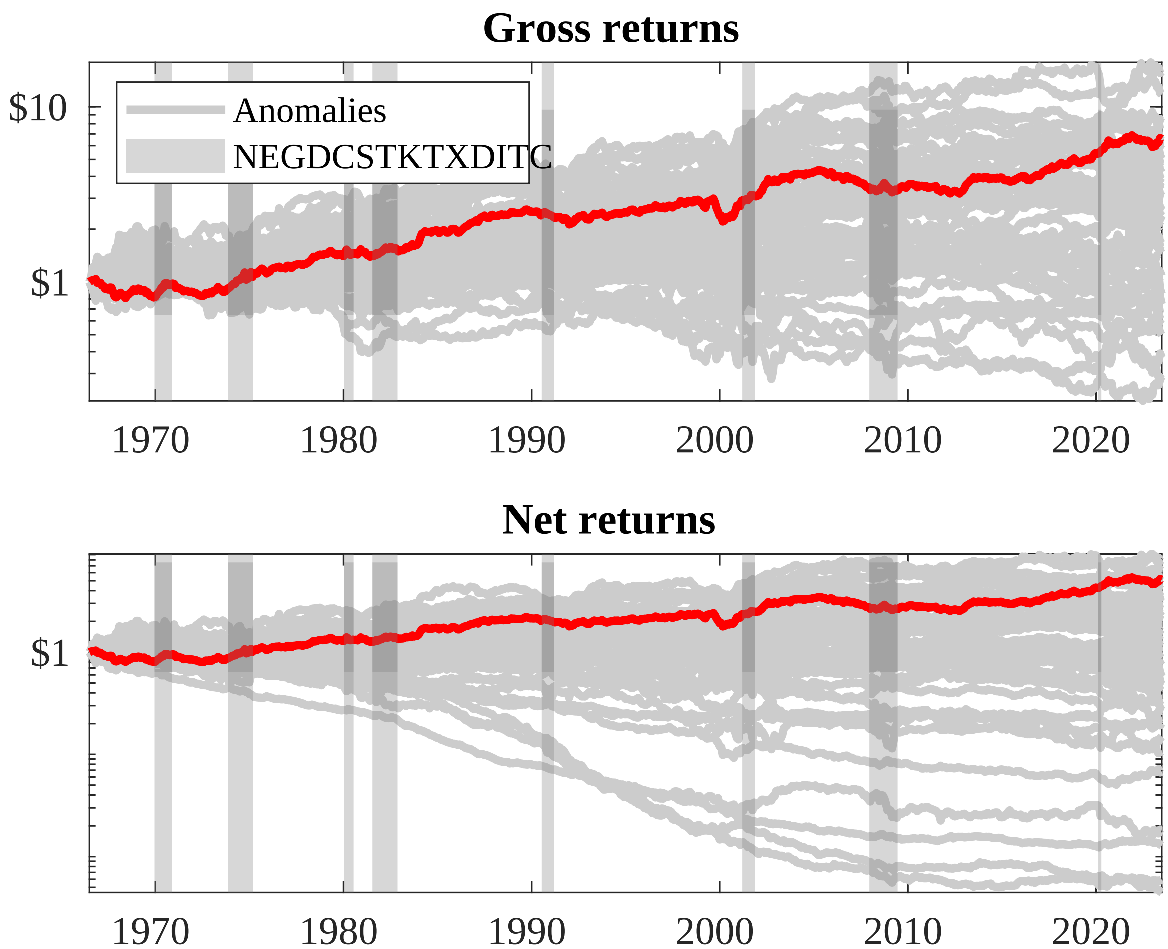


Figure 3: Dollar invested.
This figure plots the growth of a \$1 invested in 212 anomaly trading strategies (gray lines), and compares those with the TSD trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

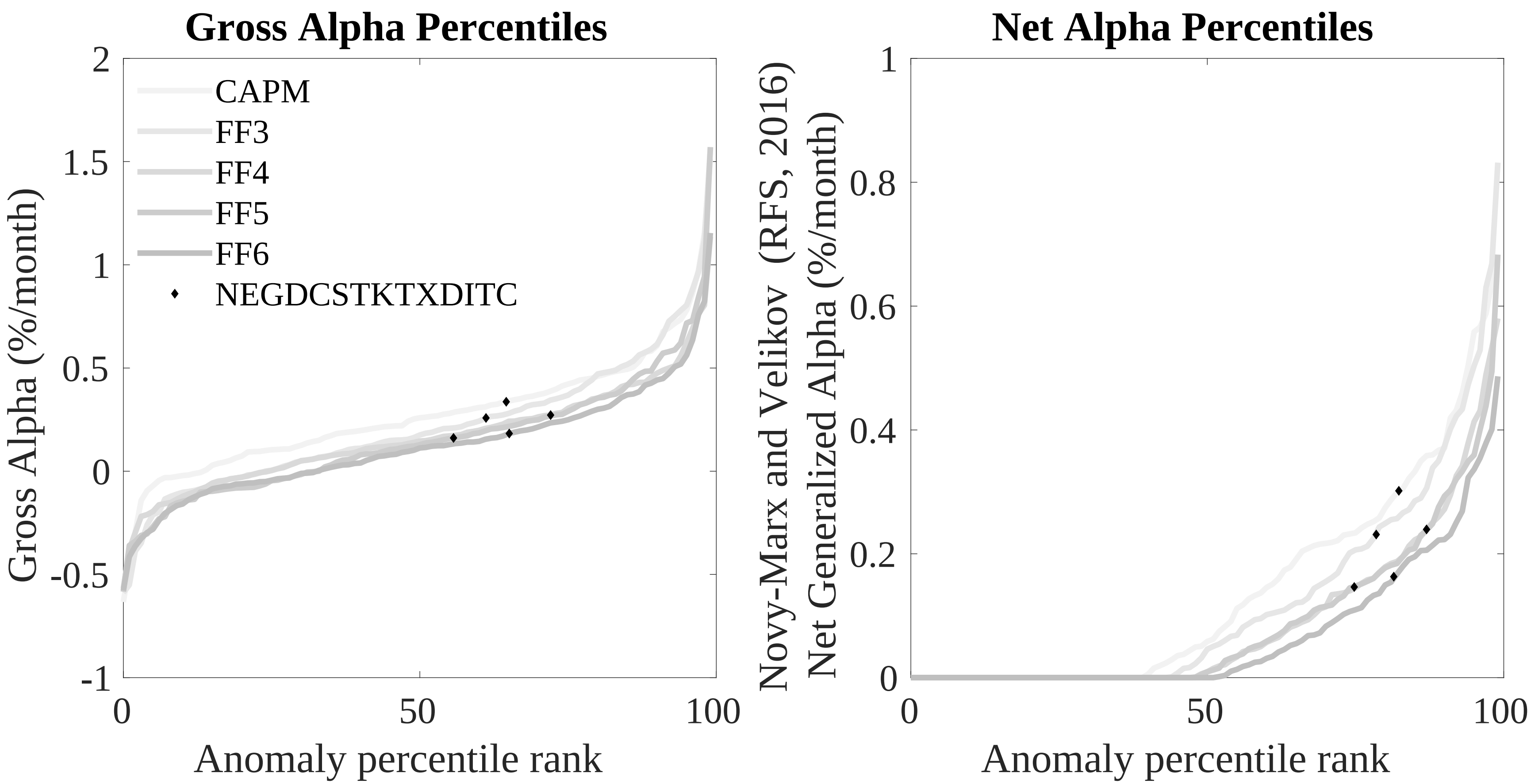
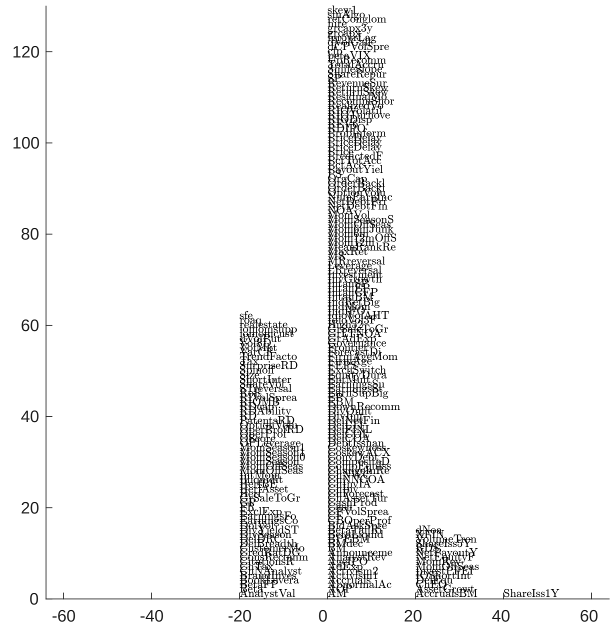
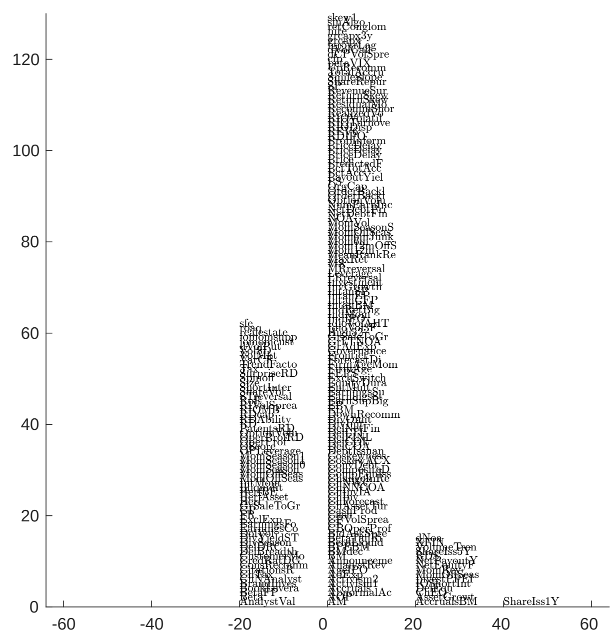


Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the TSD trading strategy alphas (diamonds). The strategies are constructed using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



Panel A: Pearson



Panel B: Spearman

Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 207 filtered anomaly signals with TSD. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

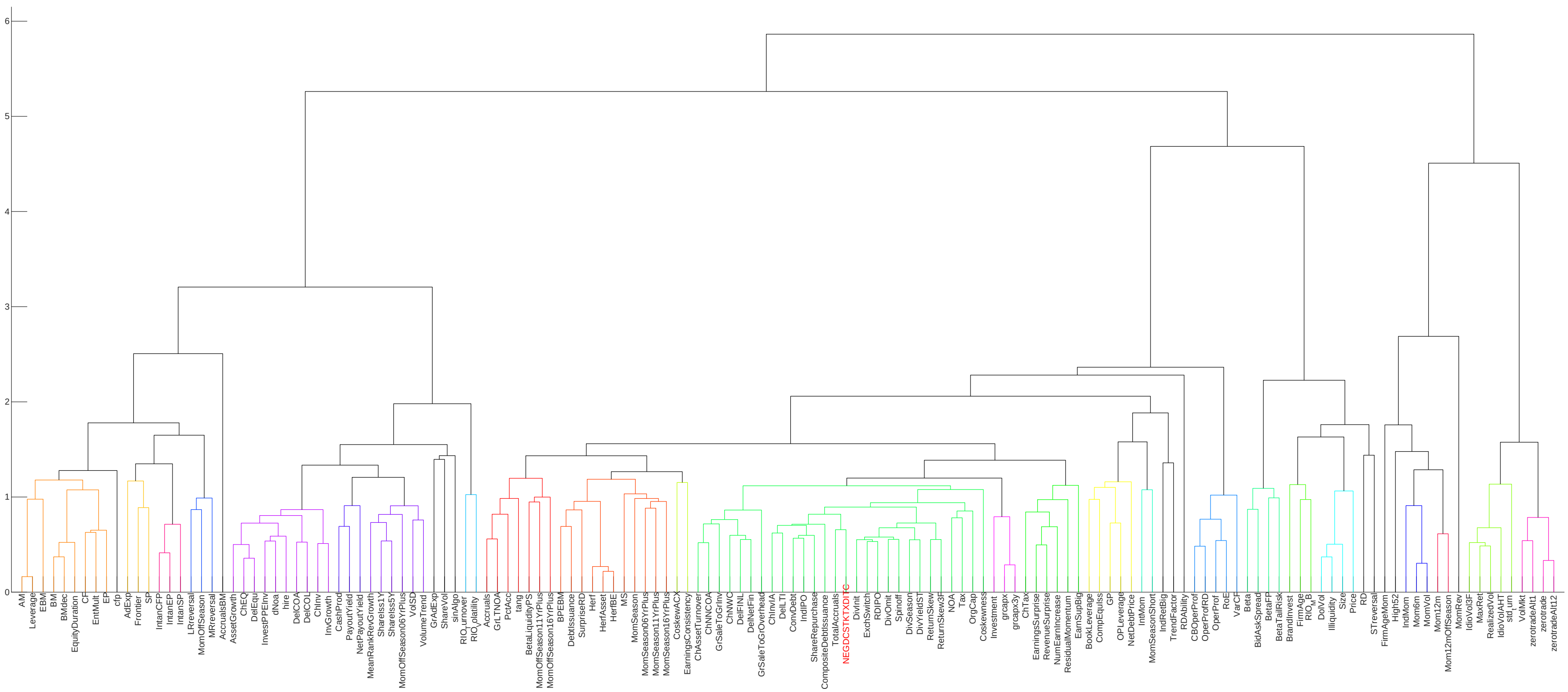


Figure 6: Agglomerative hierarchical cluster plot

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

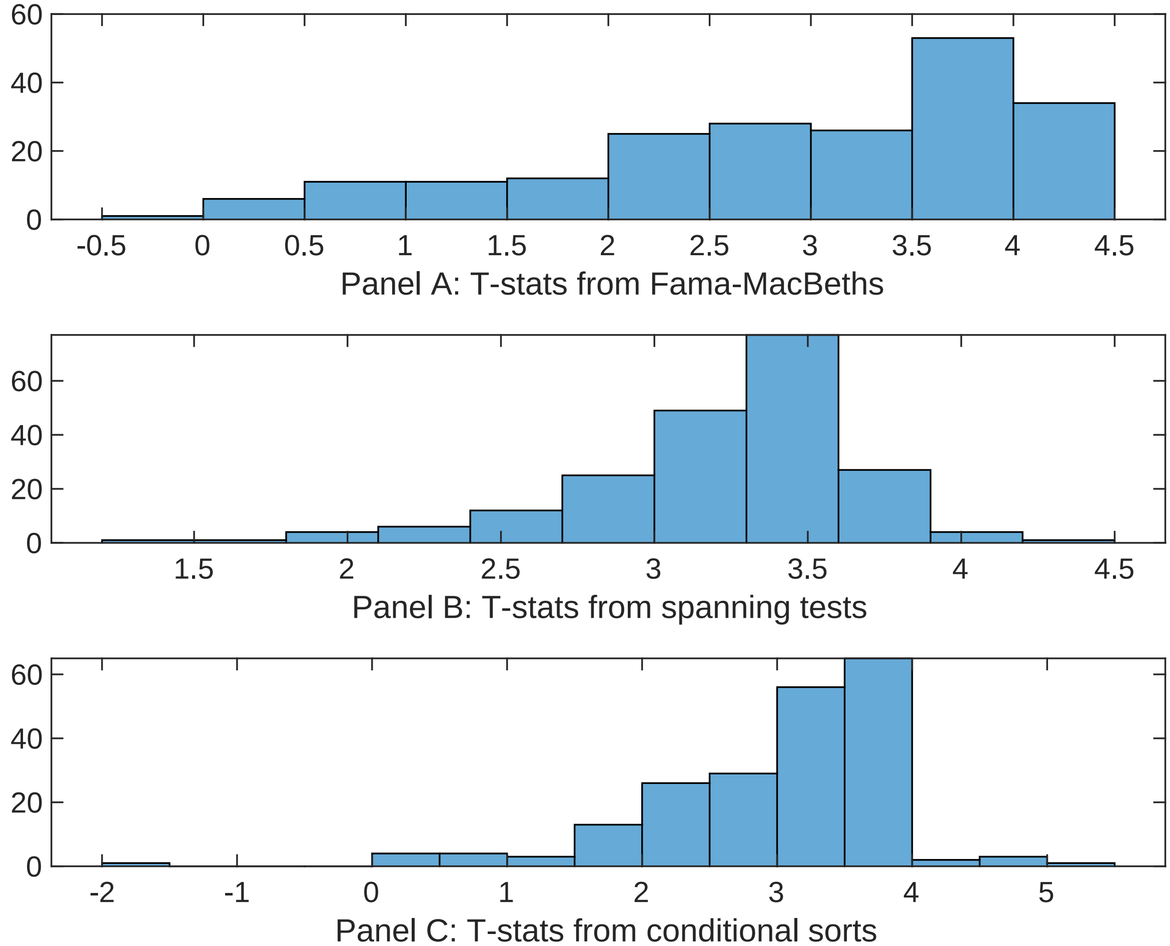


Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of TSD conditioning on each of the 207 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{TSD} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{TSD} TSD_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 207 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{TSD,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 207 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 207 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on TSD. Stocks are finally grouped into five TSD portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted TSD trading strategies conditioned on each of the 207 filtered anomalies.

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

This table presents Fama-MacBeth results of returns on TSD. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{TSD}TSD_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k}X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Net Payout Yield, Share issuance (1 year), 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	0.12	0.13	0.18	0.13	0.13	0.13	0.12
	[5.46]	[5.91]	[7.19]	[6.28]	[5.97]	[6.39]	[4.18]
TSD	0.41	0.70	0.58	0.59	0.52	0.47	0.25
	[2.09]	[3.69]	[2.93]	[2.92]	[2.73]	[2.54]	[1.14]
Anomaly 1	0.25						0.20
	[2.07]						[1.82]
Anomaly 2		0.15					0.17
		[3.28]					[0.32]
Anomaly 3			0.50				-0.76
			[3.79]				[-0.36]
Anomaly 4				0.25			0.11
				[2.35]			[1.15]
Anomaly 5					0.16		-0.29
					[3.73]		[-0.04]
Anomaly 6						1.00	0.63
						[7.72]	[5.16]
# months	679	679	684	679	684	684	679
$\bar{R}^2(\%)$	1	0	0	0	1	0	0

Table 5: Spanning tests controlling for most closely related anomalies

This table presents spanning tests results of regressing returns to the TSD trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{TSD} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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 Net Payout Yield, Share issuance (1 year), 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	0.18	0.15	0.18	0.15	0.20	0.19	0.16
	[2.20]	[1.91]	[2.25]	[1.82]	[2.48]	[2.27]	[2.00]
Anomaly 1	18.39						8.93
	[5.92]						[2.47]
Anomaly 2		23.61					13.59
		[5.75]					[2.83]
Anomaly 3			30.05				21.82
			[6.75]				[3.36]
Anomaly 4				12.74			-0.89
				[3.00]			[-0.19]
Anomaly 5					22.49		1.84
					[5.24]		[0.30]
Anomaly 6						5.60	-12.74
						[1.03]	[-2.23]
mkt	3.29	2.19	1.18	2.15	-0.16	0.21	3.35
	[1.71]	[1.16]	[0.62]	[1.09]	[-0.09]	[0.11]	[1.73]
smb	1.44	-0.95	-3.45	-2.59	-2.81	-2.96	0.70
	[0.52]	[-0.35]	[-1.26]	[-0.93]	[-1.01]	[-1.03]	[0.25]
hml	3.33	7.56	6.42	7.14	7.08	9.76	3.09
	[0.86]	[2.05]	[1.75]	[1.81]	[1.91]	[2.60]	[0.79]
rmw	0.04	2.72	11.69	8.12	12.36	10.02	2.62
	[0.01]	[0.70]	[3.18]	[2.13]	[3.30]	[2.64]	[0.60]
cma	14.54	16.93	-2.24	24.65	4.17	20.78	7.38
	[2.46]	[2.93]	[-0.32]	[4.33]	[0.59]	[2.42]	[0.87]
umd	-1.56	-3.35	-3.61	-2.95	-2.60	-3.14	-3.16
	[-0.84]	[-1.80]	[-1.93]	[-1.56]	[-1.37]	[-1.62]	[-1.70]
# months	680	680	684	680	684	684	680
$\bar{R}^2$(%)	22	21	21	18	19	15	25

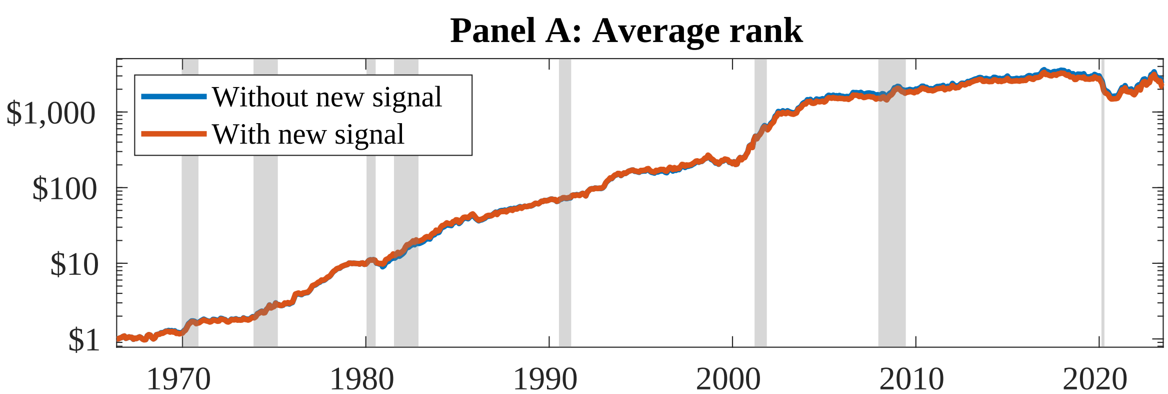


Figure 8: Combination strategy performance
This figure plots the growth of a \$1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 155 anomalies. The red solid lines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

References

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

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

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

DellaVigna, S. and Pollet, J. M. (2009). Investor inattention and friday earnings announcements. *Journal of Finance*, 64(2):709–749.

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.

Graham, J. R. and Tucker, A. L. (2006). Tax shelters and corporate debt policy. *Journal of Financial Economics*, 81(3):563–594.

Grossman, S. J. and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408.

Hanlon, M. and Heitzman, S. (2010). A review of tax research. *Journal of Accounting and Economics*, 50(2-3):127–178.

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

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*.