Tax-Effectiveness Yield and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Tax-Effectiveness Yield (TEY), 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 TEY achieves an annualized gross (net) Sharpe ratio of 0.28 (0.26), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 31 (25) bps/month with a t-statistic of 2.71 (2.17), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (Change in equity to assets, Growth in book equity, Long-term EPS forecast, Payout Yield, Asset growth, Change in current operating assets) is 37 bps/month with a t-statistic of 2.93.

1 Introduction

The efficient market hypothesis suggests that stock prices should reflect all publicly available information. However, a growing body of literature documents persistent patterns in stock returns that appear to contradict market efficiency (Harvey et al., 2016). These patterns, often called anomalies, raise important questions about market efficiency and the mechanisms through which information is incorporated into prices. While hundreds of return predictors have been documented, relatively few studies examine how firms' tax strategies affect their stock returns, despite taxation being one of the largest cash flow impacts on corporations (Graham et al., 2017). This gap is particularly notable given that tax planning activities can significantly affect firms' after-tax cash flows and risk profiles.

Prior research shows that investors face significant challenges in processing complex tax information (Plumlee and Maydew, 2014; **?**). The opacity of corporate tax strategies, combined with the technical nature of tax accounting, creates potential information frictions that could lead to systematic mispricing. These frictions may be especially pronounced given that even sophisticated market participants often struggle to fully incorporate tax-related information into their valuation models.

We propose that a firm's Tax-Effectiveness Yield (TEY) - a measure capturing the efficiency of corporate tax planning relative to peer firms - contains valuable information about future stock returns. Our hypothesis builds on three theoretical foundations. First, effective tax planning represents a form of managerial skill that may signal broader operational excellence (Dyreng and Hanlon, 2010). Managers who successfully navigate complex tax environments likely possess superior information processing and risk management capabilities that benefit shareholders across multiple dimensions.

Second, tax effectiveness can create persistent competitive advantages through lower effective tax rates and enhanced after-tax cash flows (**?**). These advantages may be particularly valuable during periods of regulatory change or economic stress when tax planning flexibility provides strategic options. The sustainability of tax advantages suggests their value may not be fully reflected in current stock prices.

Third, information processing frictions in tax-related disclosures can lead to systematic underreaction to tax planning signals (?). The technical nature of tax accounting, combined with the delayed revelation of tax strategy outcomes, creates conditions where sophisticated tax planning may be undervalued by the market. This builds on theoretical work showing that complex information requiring specialized expertise is often slowly incorporated into prices (Hong and Stein, 1999).

Our empirical analysis reveals that TEY strongly predicts future stock returns. A value-weighted long-short portfolio strategy based on TEY quintiles generates monthly abnormal returns of 31 basis points relative to the Fama-French five-factor model plus momentum (t-statistic = 2.71). The strategy's economic significance is substantial, with an annualized gross Sharpe ratio of 0.28. These results are robust to controlling for transaction costs, with net returns remaining both economically and statistically significant at 25 basis points per month (t-statistic = 2.17).

Importantly, TEY's predictive power persists after controlling for known return predictors. When we account for the six most closely related anomalies from the literature, including changes in equity-to-assets and payout yield, the strategy's alpha increases to 37 basis points monthly (t-statistic = 2.93). This suggests that TEY captures unique information not contained in existing factors or anomalies.

The signal's robustness is further demonstrated through various portfolio construction approaches and subsamples. TEY maintains significant predictive power across different size quintiles, though its effect is strongest among larger, more liquid stocks where sophisticated tax planning is more prevalent. This pattern distinguishes TEY from many anomalies that are concentrated in small, illiquid stocks.

Our study makes several important contributions to the asset pricing and corporate finance literatures. First, we introduce a novel return predictor that bridges the gap between tax accounting research and asset pricing. While prior work has examined how specific tax events affect stock prices (Hanlon and Heitzman, 2010), we show that systematic differences in firms' tax planning effectiveness contain valuable information about future returns. This extends our understanding of how corporate tax strategies affect firm value.

Second, we contribute to the growing literature on information processing frictions in financial markets (Hirshleifer and Teoh, 2003). Our findings suggest that the complexity of tax information creates persistent mispricing that sophisticated investors can exploit. This adds to evidence that technical information requiring specialized expertise is often slowly incorporated into stock prices (**?**).

Finally, our results have important implications for both academic research and practice. For researchers, we demonstrate the value of incorporating tax planning effectiveness into asset pricing models. For practitioners, our findings suggest that systematic analysis of corporate tax strategies can identify profitable trading opportunities. More broadly, our work highlights how the intersection of tax policy and market efficiency creates opportunities for both academic inquiry and investment strategy development.

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-Effectiveness Yield, which is constructed as the ratio of tax credits to dividend payments. 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 TXC for tax credits and item DVC for dividend payments. Tax credits (TXC) represent the firm's tax benefits or credits received from various sources, including investment incentives, research and development activities, and other tax-advantaged operations. Dividend payments (DVC), on the other hand, reflect the total amount of cash dividends declared on common stock during the fiscal year, representing a direct form of shareholder return. The construction of the signal follows a straightforward ratio format, where we divide TXC by DVC for each firm in each year of our sample. This ratio captures the relative scale of a firm's tax benefits against its dividend distributions, offering insight into how effectively the firm leverages tax advantages while maintaining shareholder distributions. By focusing on this relationship, the signal aims to reflect aspects of tax efficiency and shareholder return management in a manner that is both scalable and interpretable. We construct this ratio using end-of-fiscal-year values for both TXC and DVC to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the TEY signal. Panel A plots the time-series of the mean, median, and interquartile range for TEY. On average, the cross-sectional mean (median) TEY is 5.47 (1.39) over the 1971 to 2023 sample, where the starting date is determined by the availability of the input TEY data. The signal's interquartile range spans 0.07 to 4.58. Panel B of Figure 1 plots the time-series of the coverage of the TEY signal for the CRSP universe. On average, the TEY signal is available for 2.39% of CRSP names, which on average make up 5.62% of total market capitalization.

4 Does TEY predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on TEY 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 TEY portfolio and sells the low TEY 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 TEY strategy earns an average return of 0.29% per month with a t-statistic of 2.03. The annualized Sharpe ratio of the strategy is 0.28. The alphas range from 0.12% to 0.33% per month and have t-statistics exceeding 0.92 everywhere. The lowest alpha is with respect to the CAPM 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 -0.58, with a t-statistic of -7.48 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 233 stocks and an average market capitalization of at least $748 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 NYSE breakpoints and equal-weighted portfolios, and equals 22 bps/month with a t-statistics of 2.44. Out of the twenty-five alphas reported in Panel A, the t-statistics for eighteen exceed two, and for five 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 11-48bps/month. The lowest return, (11 bps/month), is achieved from the quintile sort using NYSE breakpoints and equal-weighted portfolios, and has an associated t-statistic of 1.25. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the TEY trading strategy improves the achievable mean-variance efficient frontier spanned by the factor models in twenty-three cases, and significantly expands the achievable frontier in twelve cases.

Table 3 provides direct tests for the role size plays in the TEY strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and TEY, as well as average returns and alphas for long/short trading TEY 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 TEY strategy achieves an average return of 22 bps/month with a t-statistic of 1.45. Among these large cap stocks, the alphas for the TEY strategy relative to the five most common factor models range from 3 to 25 bps/month with t-statistics between 0.24 and 1.92.

5 How does TEY perform relative to the zoo?

Figure 2 puts the performance of TEY 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 TEY strategy falls in the distribution. The TEY strategy's gross (net) Sharpe ratio of 0.28 (0.26) is greater than 59% (85%) 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 TEY strategy (red line).[2] Ignoring trading costs, a $1 invested in the TEY strategy would have yielded $3.11 which ranks the TEY strategy in the top 6% across the 212 anomalies. Accounting for trading costs, a $1 invested in the TEY strategy would have yielded $2.52 which ranks the TEY strategy in the top 5% across the 212 anomalies.

Figure 4 plots percentile ranks for the 212 anomaly trading strategies in terms

[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.

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 TEY relative to those. Panel A shows that the TEY strategy gross alphas fall between the 30 and 82 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 197106 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 TEY strategy has a positive net generalized alpha for five out of the five factor models. In these cases TEY ranks between the 54 and 92 percentiles in terms of how much it could have expanded the achievable investment frontier.

6 Does TEY 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 TEY with 202 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 TEY or at least to weaken the power TEY has predicting the cross-section of returns. Figure 7 plots histograms

[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.

of t-statistics for predictability tests of TEY conditioning on each of the 202 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{TEY} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{TEY}TEY_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 202 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{TEY,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 202 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 202 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on TEY. Stocks are finally grouped into five TEY portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted TEY trading strategies conditioned on each of the 202 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on TEY 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 TEY signal in these Fama-MacBeth regressions exceed 0.65, with the minimum t-statistic occurring when controlling for Payout Yield. Controlling for all six closely related anomalies, the t-statistic on TEY is 1.71.

Similarly, Table 5 reports results from spanning tests that regress returns to the TEY 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 TEY strategy earns alphas that range from 29-38bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.40, which is achieved when controlling for Payout Yield. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the TEY trading strategy achieves an alpha of 37bps/month with a t-statistic of 2.93.

7 Does TEY add relative to the whole zoo?

Finally, we can ask how much adding TEY 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 156 anomalies from the zoo that satisfy our inclusion criteria (blue lines) or these 156 anomalies augmented with the TEY 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 156-anomaly combination strategy grows to $1203.74, while $1 investment in the combination strategy that includes TEY grows to $1041.48.

8 Conclusion

This study provides compelling evidence for the predictive power of Tax-Effectiveness Yield (TEY) in forecasting cross-sectional stock returns. Our findings demonstrate that TEY generates economically and statistically significant returns, with a value-

[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 TEY is available.

weighted long/short strategy achieving an impressive annualized Sharpe ratio of 0.28 (0.26 net of transaction costs). The strategy's robustness is particularly noteworthy, maintaining significant abnormal returns of 31 basis points per month (25 bps net) even after controlling for the Fama-French five factors and momentum. Moreover, the signal's predictive power persists when controlling for six closely related factors from the factor zoo, yielding a substantial alpha of 37 bps per month.

These results have important implications for both academic research and investment practice. For academics, our findings contribute to the growing literature on return predictability and suggest that tax considerations play a meaningful role in asset pricing. For practitioners, TEY appears to be a valuable tool for portfolio construction and alpha generation, maintaining its effectiveness even after accounting for transaction costs.

However, several limitations should be noted. First, our analysis focuses on U.S. equity markets, and the signal's effectiveness in international markets remains to be tested. Second, the study period may not fully capture the impact of recent changes in tax regulations and market structure.

Future research could explore the interaction between TEY and other established factors, investigate its performance in different market regimes, and examine its applicability across different asset classes and geographical markets. Additionally, researchers might investigate the underlying economic mechanisms driving the TEY premium and its potential variation across different tax jurisdictions.

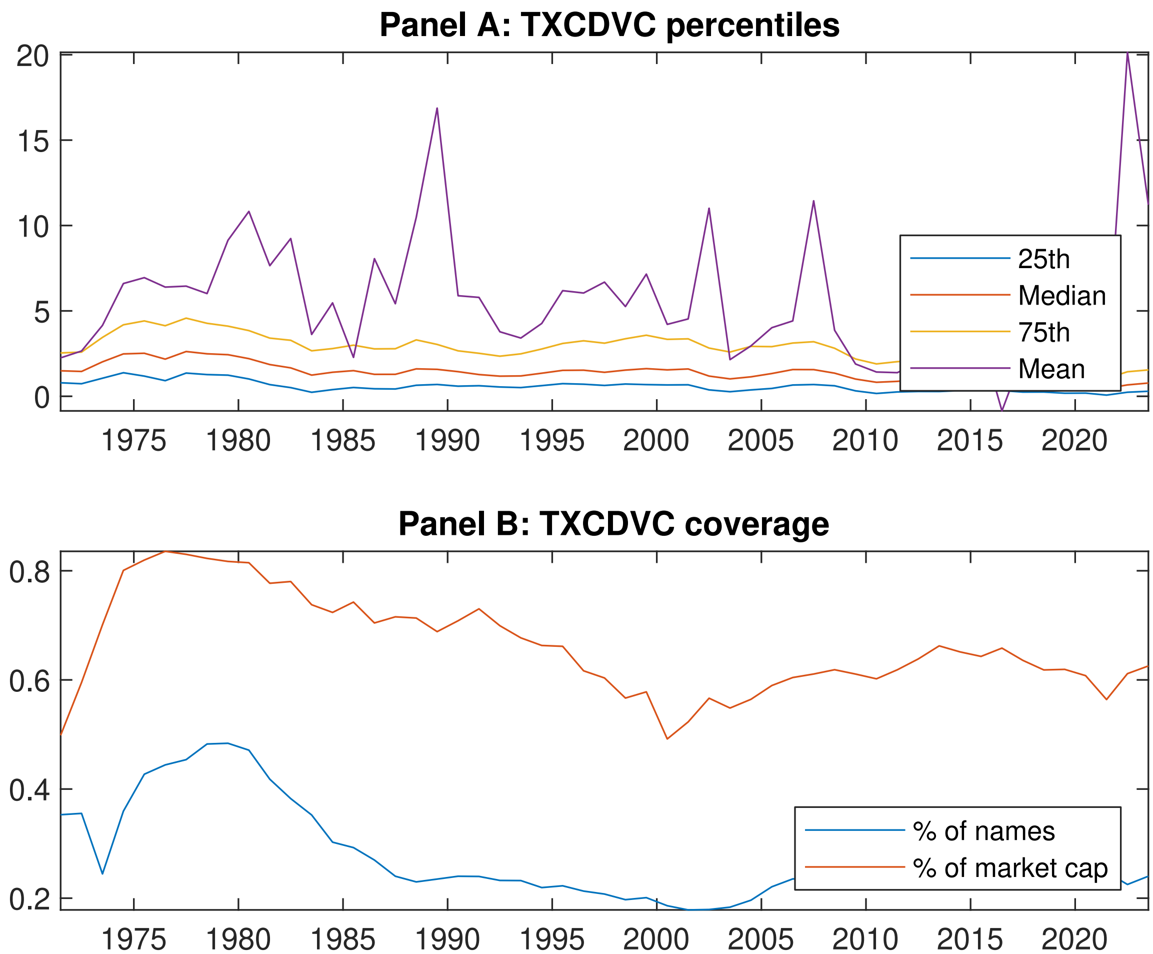


Figure 1: Times series of TEY percentiles and coverage.
This figure plots descriptive statistics for TEY. Panel A shows cross-sectional percentiles of TEY over the sample. Panel B plots the monthly coverage of TEY relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on TEY. 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 197106 to 202306.

Panel A: Excess returns and alphas on TEY-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.51	0.56	0.60	0.68	0.80	0.29
	[2.88]	[3.43]	[3.45]	[3.53]	[3.72]	[2.03]
α_{CAPM}	0.02	0.08	0.08	0.07	0.14	0.12
	[0.17]	[1.11]	[1.13]	[1.36]	[1.88]	[0.92]
α_{FF3}	-0.14	0.03	0.04	0.08	0.17	0.30
	[-1.79]	[0.43]	[0.57]	[1.44]	[2.25]	[2.53]
α_{FF4}	-0.13	0.00	0.04	0.10	0.17	0.30
	[-1.62]	[0.01]	[0.66]	[1.76]	[2.31]	[2.46]
α_{FF5}	-0.20	-0.15	-0.17	-0.01	0.13	0.33
	[-2.67]	[-2.65]	[-2.90]	[-0.28]	[1.81]	[2.89]
α_{FF6}	-0.18	-0.15	-0.15	0.01	0.13	0.31
	[-2.37]	[-2.67]	[-2.51]	[0.14]	[1.85]	[2.71]
Panel B: Fama and French (2018) 6-factor model loadings for TEY-sorted portfolios						
β_{MKT}	0.93	0.92	0.96	1.04	1.06	0.14
	[52.09]	[69.69]	[69.92]	[83.13]	[63.01]	[5.13]
β_{SMB}	-0.11	-0.21	-0.07	-0.04	0.13	0.24
	[-3.97]	[-10.68]	[-3.27]	[-1.97]	[5.15]	[5.83]
β_{HML}	0.26	0.01	0.01	-0.06	-0.02	-0.28
	[7.59]	[0.57]	[0.27]	[-2.55]	[-0.75]	[-5.46]
β_{RMW}	-0.06	0.22	0.43	0.21	0.22	0.28
	[-1.75]	[8.69]	[16.13]	[8.65]	[6.74]	[5.36]
β_{CMA}	0.39	0.39	0.24	0.09	-0.19	-0.58
	[7.65]	[10.21]	[5.98]	[2.59]	[-3.91]	[-7.48]
β_{UMD}	-0.03	0.00	-0.03	-0.03	-0.01	0.02
	[-1.75]	[0.35]	[-2.40]	[-2.69]	[-0.38]	[0.91]
Panel C: Average number of firms (n) and market capitalization (me)						
n	261	233	242	267	331	
me ($\$10^6$)	748	1693	1739	1595	1097	

Table 2: Robustness to sorting methodology & trading costs

This table evaluates the robustness of the choices made in the TEY 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 197106 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	0.29	0.12	0.30	0.30	0.33	0.31
			[2.03]	[0.92]	[2.53]	[2.46]	[2.89]	[2.71]
Quintile	NYSE	EW	0.22	0.12	0.19	0.17	0.10	0.09
			[2.44]	[1.42]	[2.28]	[2.06]	[1.37]	[1.18]
Quintile	Name	VW	0.32	0.15	0.33	0.34	0.36	0.35
			[2.15]	[1.05]	[2.68]	[2.65]	[3.04]	[2.90]
Quintile	Cap	VW	0.23	0.06	0.21	0.24	0.27	0.28
			[1.73]	[0.45]	[1.87]	[2.14]	[2.49]	[2.59]
Decile	NYSE	VW	0.51	0.36	0.55	0.57	0.58	0.59
			[3.08]	[2.24]	[3.67]	[3.73]	[3.97]	[3.93]

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	0.10	0.26	0.25	0.27	0.25
			[1.86]	[0.74]	[2.12]	[2.11]	[2.33]	[2.17]
Quintile	NYSE	EW	0.11	0.02	0.07	0.07		
			[1.25]	[0.21]	[0.87]	[0.79]		
Quintile	Name	VW	0.29	0.12	0.28	0.29	0.30	0.27
			[1.98]	[0.86]	[2.25]	[2.27]	[2.44]	[2.31]
Quintile	Cap	VW	0.21	0.03	0.16	0.18	0.20	0.20
			[1.58]	[0.21]	[1.43]	[1.62]	[1.83]	[1.82]
Decile	NYSE	VW	0.48	0.32	0.48	0.49	0.50	0.48
			[2.87]	[1.96]	[3.20]	[3.26]	[3.36]	[3.30]

Table 3: Conditional sort on size and TEY

This table presents results for conditional double sorts on size and TEY. 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 TEY. 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 TEY and short stocks with low TEY .Panel B documents the average number of firms and the average firm size for each portfolio. The sample period is 197106 to 202306.

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

		TEY Quintiles					TEY Strategies					
		(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Size quintiles	(1)	0.74	1.02	0.86	1.31	1.01	0.27	0.19	0.24	0.16	0.20	0.13
		[3.05]	[4.09]	[3.97]	[3.32]	[3.92]	[1.85]	[1.35]	[1.70]	[1.14]	[1.37]	[0.91]
	(2)	0.80	0.80	0.83	0.87	0.87	0.07	-0.03	0.03	0.04	-0.08	-0.07
		[3.61]	[3.95]	[3.84]	[3.74]	[3.49]	[0.53]	[-0.21]	[0.21]	[0.35]	[-0.70]	[-0.57]
	(3)	0.69	0.69	0.87	0.92	0.88	0.19	0.06	0.15	0.16	0.04	0.05
		[3.40]	[3.65]	[4.23]	[4.14]	[3.65]	[1.41]	[0.45]	[1.21]	[1.28]	[0.39]	[0.45]
	(4)	0.64	0.74	0.77	0.72	0.77	0.14	-0.04	0.08	0.11	-0.02	0.01
		[3.29]	[3.97]	[3.89]	[3.51]	[3.25]	[0.96]	[-0.32]	[0.60]	[0.84]	[-0.12]	[0.08]
	(5)	0.50	0.56	0.57	0.63	0.72	0.22	0.03	0.18	0.19	0.25	0.24
		[3.07]	[3.36]	[3.16]	[3.29]	[3.37]	[1.45]	[0.24]	[1.34]	[1.38]	[1.92]	[1.83]

Panel B: Portfolio average number of firms and market capitalization

		TEY Quintiles Average n					TEY Quintiles Average market capitalization ($\$10^6$)				
		(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
Size quintiles	(1)	91	92	92	92	92	6	7	8	8	9
	(2)	47	47	47	47	47	20	21	21	21	21
	(3)	42	42	42	42	42	46	45	46	46	46
	(4)	41	41	41	41	41	121	124	122	122	120
	(5)	45	45	45	45	45	942	1297	1241	1369	1043

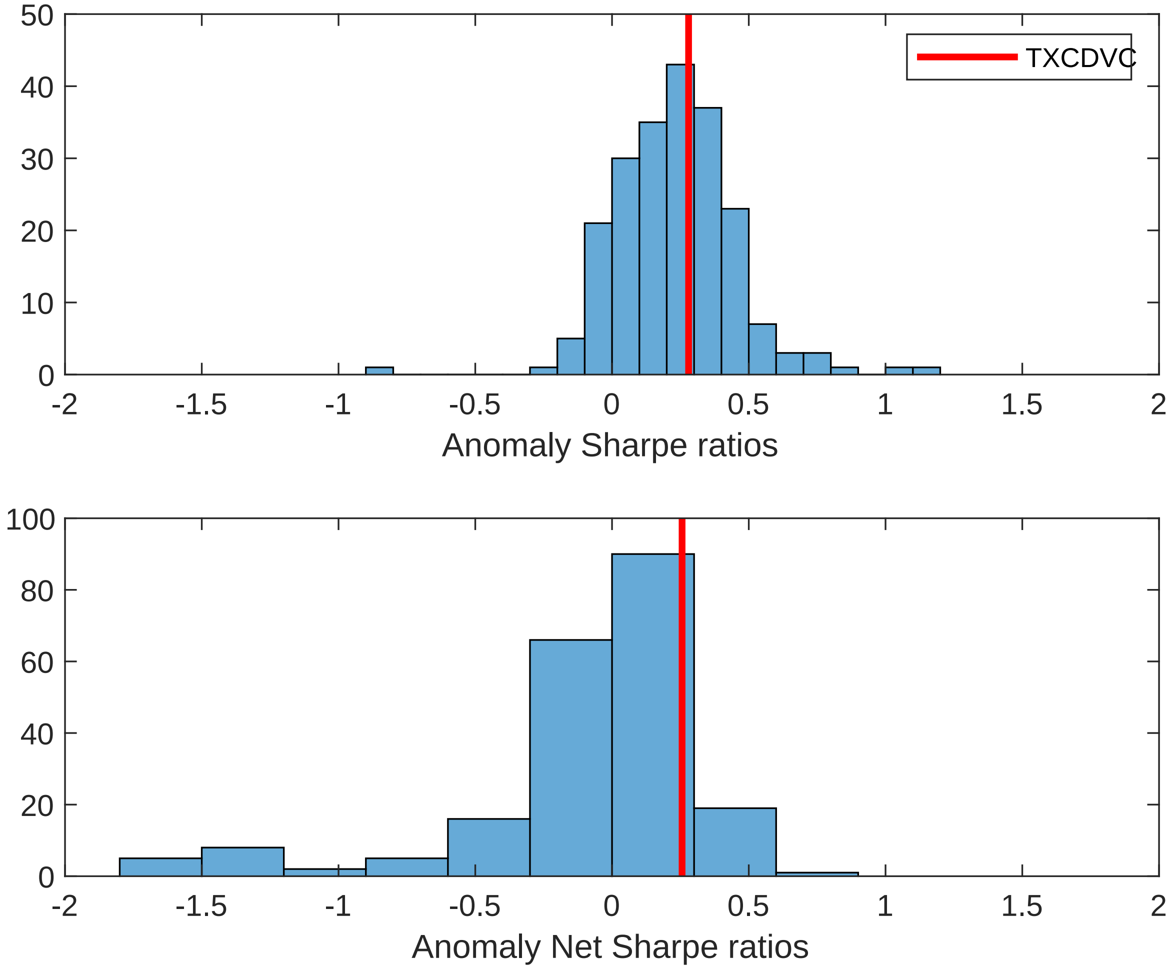


Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the TEY with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

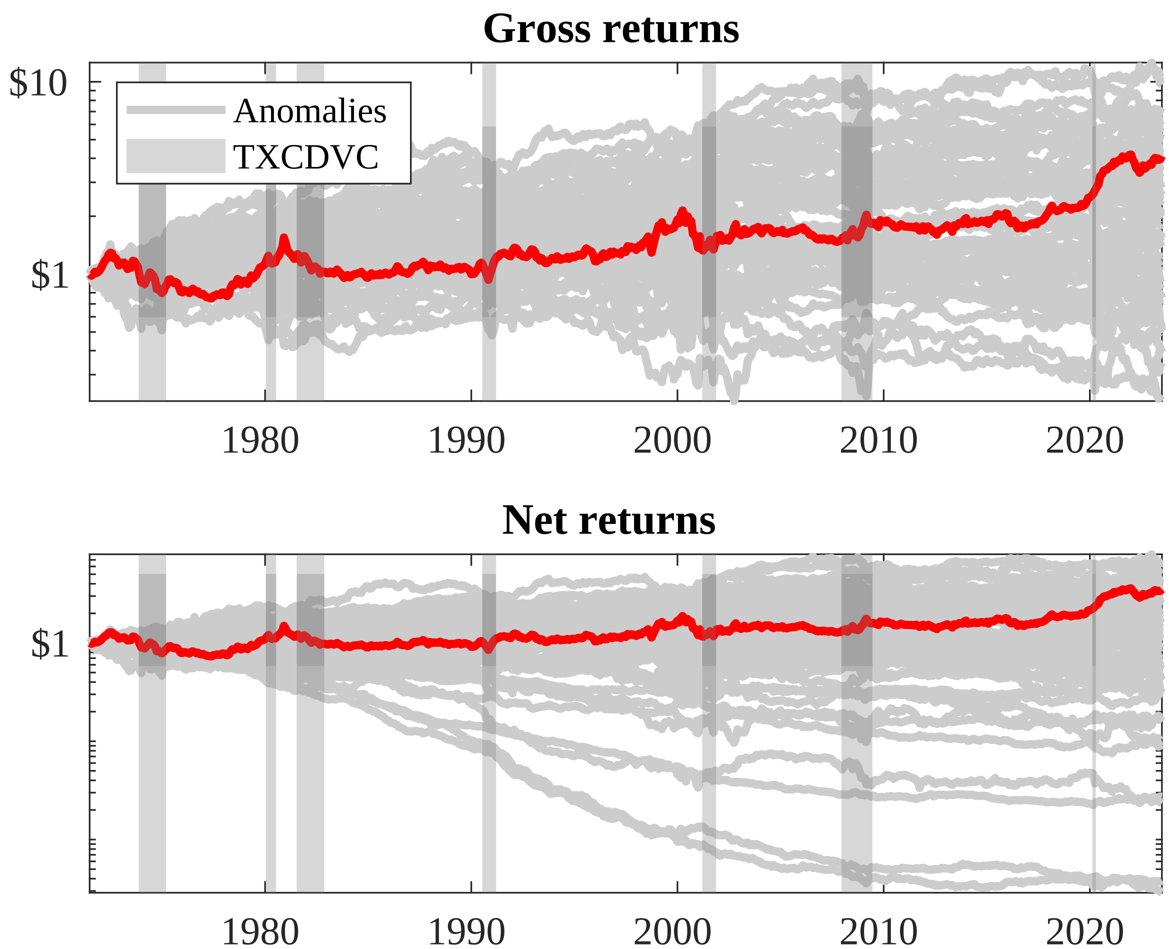


Figure 3: Dollar invested.
This figure plots the growth of a \$1 invested in 212 anomaly trading strategies (gray lines), and compares those with the TEY trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

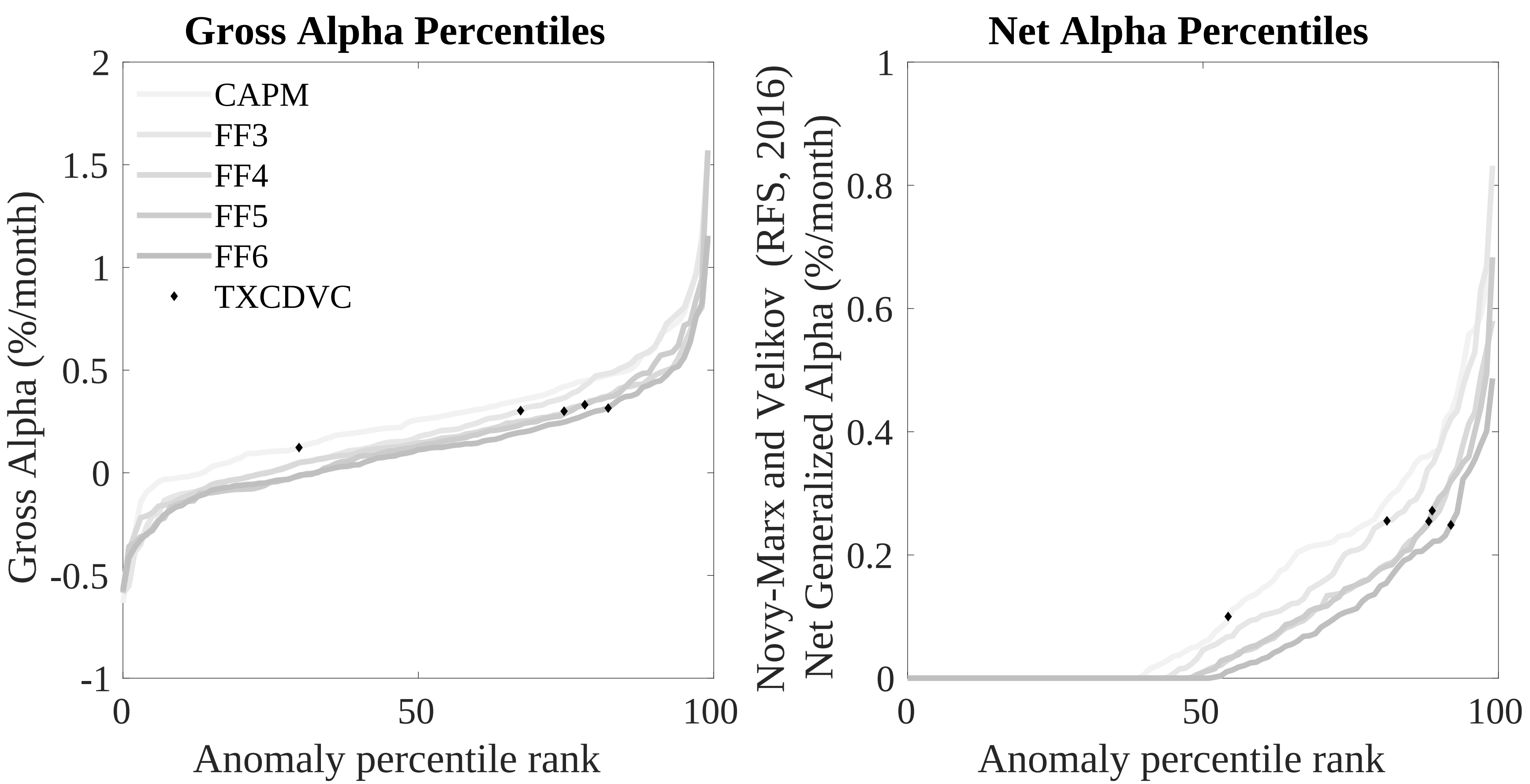
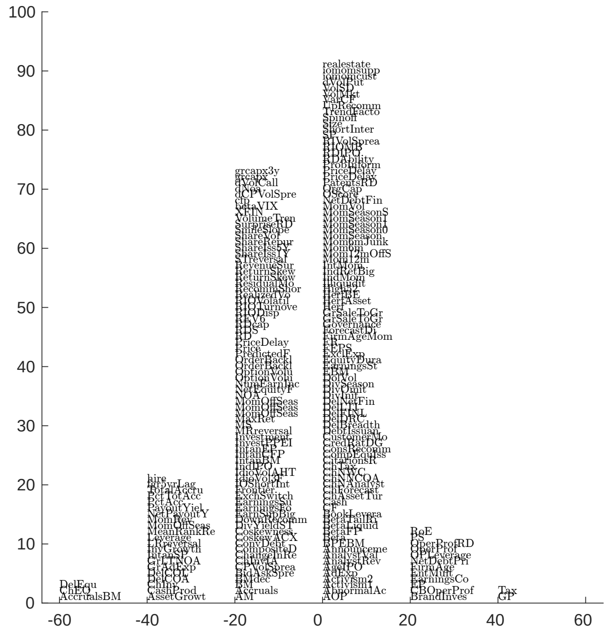
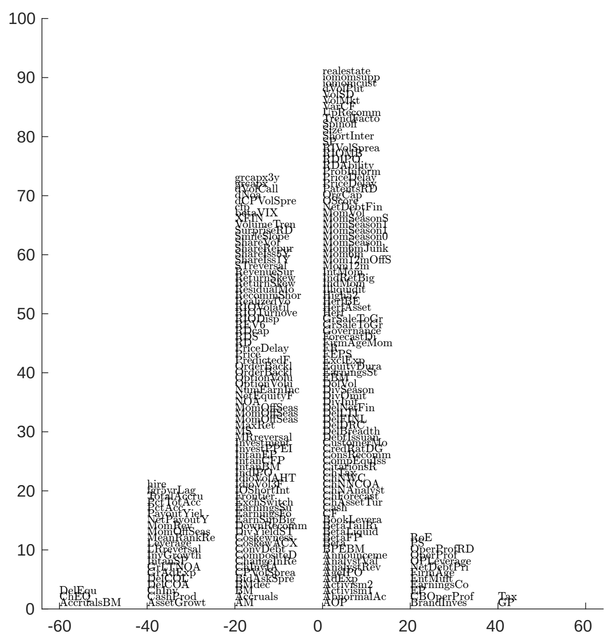


Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the TEY trading strategy alphas (diamonds). The strategies are constructed using 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 202 filtered anomaly signals with TEY. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

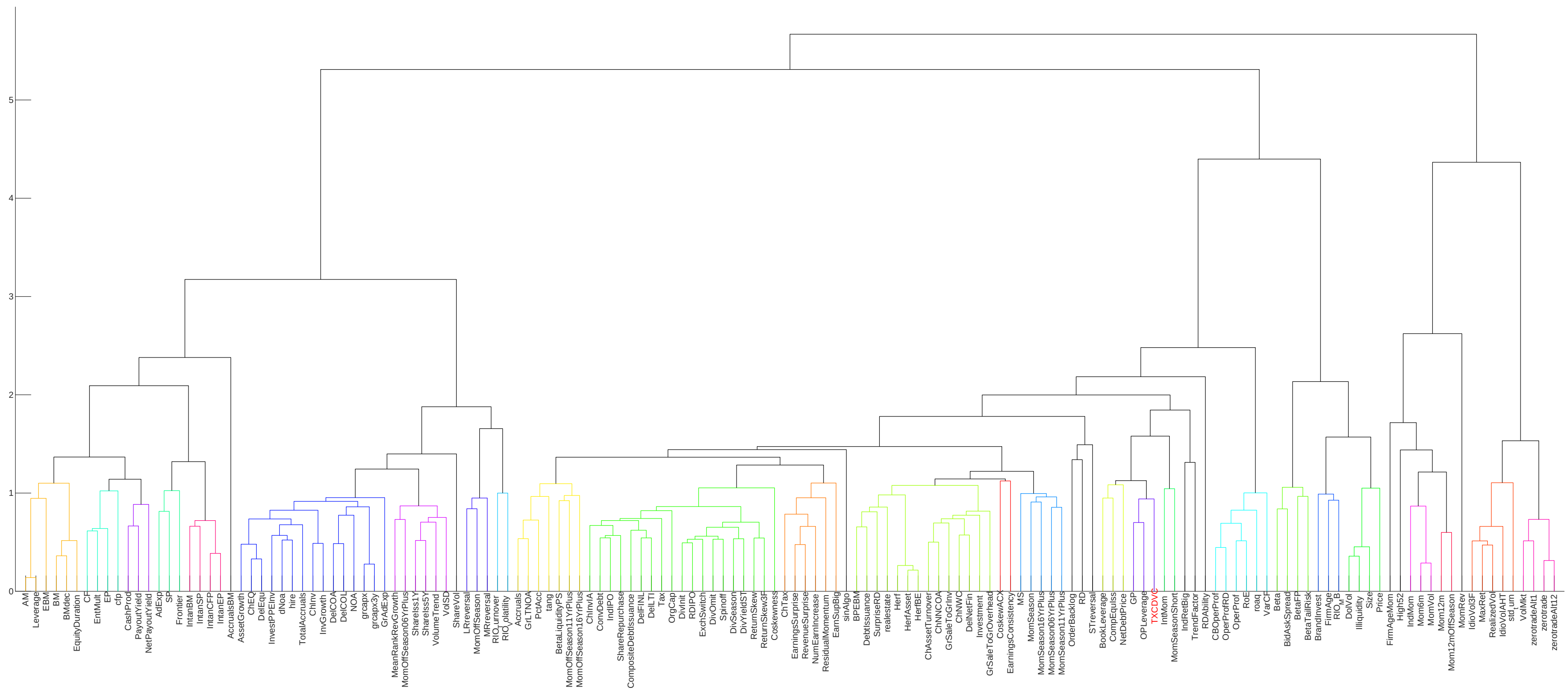


Figure 6: Agglomerative hierarchical cluster plot
This figure plots an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

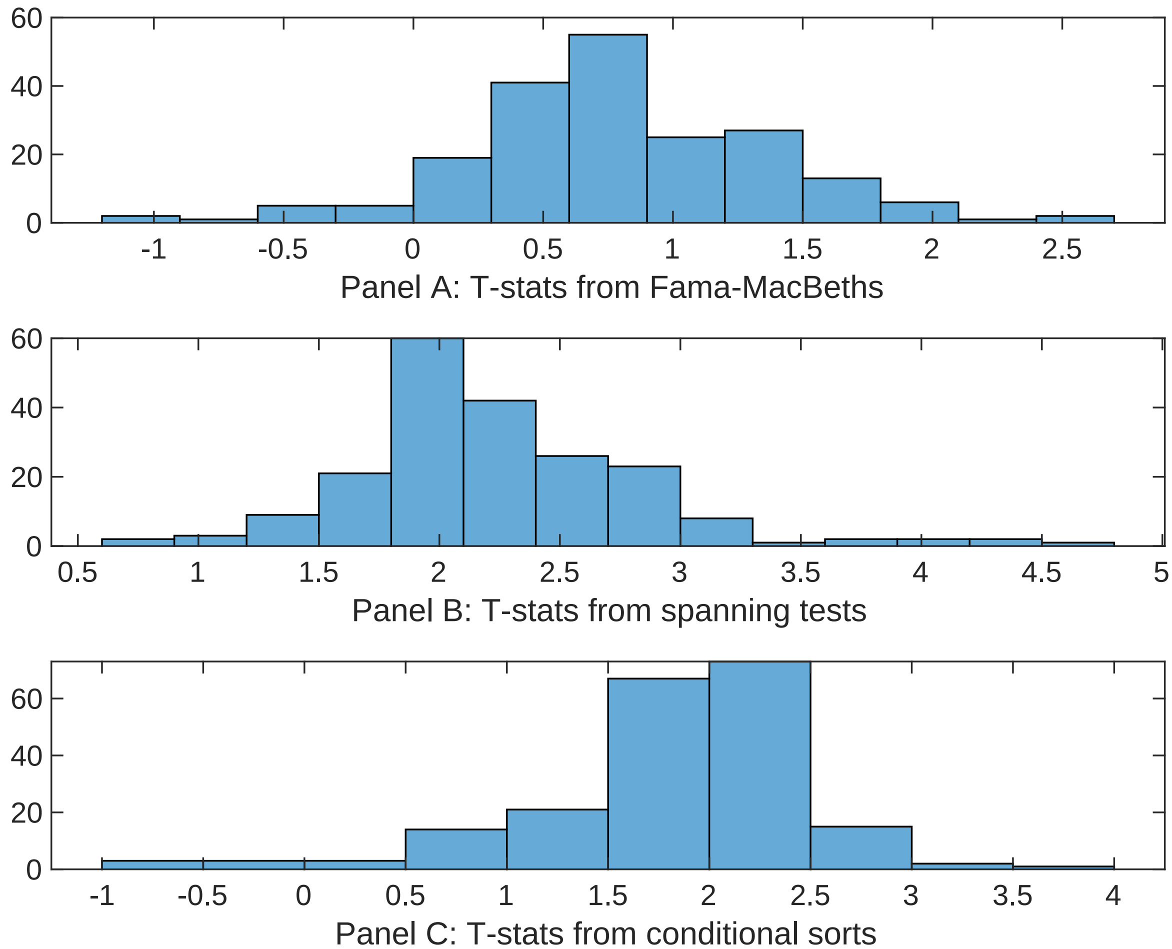


Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of TEY conditioning on each of the 202 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{TEY} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{TEY} TEY_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 202 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{TEY,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 202 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 202 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on TEY. Stocks are finally grouped into five TEY portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted TEY trading strategies conditioned on each of the 202 filtered anomalies.

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

This table presents Fama-MacBeth results of returns on TEY. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{TEY} TEY_{i,t} + \sum_{k=1}^{s} ix \beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are Change in equity to assets, Growth in book equity, Long-term EPS forecast, Payout Yield, Asset growth, Change in current operating assets. 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 197106 to 202306.

Intercept	0.13	0.19	0.13	0.12	0.13	0.13	0.14
	[6.46]	[7.76]	[7.12]	[6.25]	[6.80]	[6.52]	[6.84]
TEY	0.13	0.11	0.13	0.53	0.10	0.90	0.19
	[1.84]	[1.43]	[1.34]	[0.65]	[1.36]	[1.23]	[1.71]
Anomaly 1	0.16						-0.17
	[3.92]						[-0.30]
Anomaly 2		0.61					0.15
		[4.51]					[1.04]
Anomaly 3			0.96				0.39
			[1.26]				[0.47]
Anomaly 4				0.19			-0.16
				[0.40]			[-0.24]
Anomaly 5					0.84		0.37
					[6.88]		[2.70]
Anomaly 6						0.18	0.19
						[5.13]	[0.37]
# months	624	624	492	619	624	624	487
$\bar{R}^2(\%)$	1	1	1	1	1	1	0

Table 5: Spanning tests controlling for most closely related anomalies

This table presents spanning tests results of regressing returns to the TEY trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{TEY} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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 Change in equity to assets, Growth in book equity, Long-term EPS forecast, Payout Yield, Asset growth, Change in current operating assets. 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 197106 to 202306.

Intercept	0.29	0.33	0.31	0.38	0.29	0.31	0.37
	[2.60]	[2.93]	[2.40]	[3.37]	[2.58]	[2.63]	[2.93]
Anomaly 1	-50.16						-19.14
	[-8.48]						[-1.47]
Anomaly 2		-48.71					-15.76
		[-7.88]					[-1.29]
Anomaly 3			-36.26				-25.18
			[-8.34]				[-5.13]
Anomaly 4				-31.83			-9.92
				[-7.80]			[-1.79]
Anomaly 5					-45.33		-9.78
					[-6.07]		[-1.04]
Anomaly 6						-11.32	7.78
						[-1.92]	[1.18]
mkt	13.98	12.23	-1.80	7.40	13.70	14.04	-1.53
	[5.42]	[4.69]	[-0.54]	[2.72]	[5.17]	[5.17]	[-0.46]
smb	23.64	25.02	6.33	18.79	27.31	21.03	9.56
	[6.11]	[6.42]	[1.33]	[4.78]	[6.79]	[4.97]	[1.93]
hml	-21.78	-22.20	-4.85	-12.74	-26.11	-23.43	-7.13
	[-4.40]	[-4.45]	[-0.81]	[-2.36]	[-5.18]	[-4.11]	[-1.11]
rmw	23.69	26.19	26.02	24.58	27.98	26.98	22.27
	[4.68]	[5.16]	[4.39]	[4.83]	[5.42]	[5.04]	[3.73]
cma	-6.79	-10.82	-30.13	-44.46	-2.24	-52.72	6.90
	[-0.71]	[-1.12]	[-3.36]	[-5.68]	[-0.19]	[-6.27]	[0.54]
umd	1.47	3.40	2.96	-2.87	0.97	2.12	0.77
	[0.58]	[1.32]	[1.02]	[-1.08]	[0.37]	[0.79]	[0.26]
# months	624	624	492	620	624	624	488
$\bar{R}^2(\%)$	47	46	43	47	44	41	47

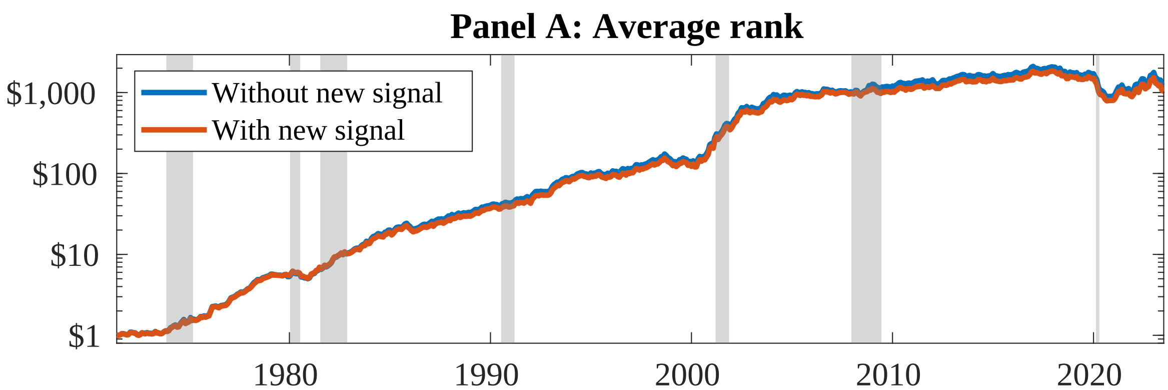


Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 156 anomalies. The red solid lines indicate combination trading strategies that utilize the 156 anomalies as well as TEY. 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.

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

Dyreng, S. D. and Hanlon, M. (2010). Long-run corporate tax avoidance. *The Accounting Review*, 85(4):1163–1189.

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., Hanlon, M., Shevlin, T., and Shroff, N. (2017). Tax rates and corporate decision-making. *Review of Financial Studies*, 30(9):3128–3175.

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

Harvey, C. R., Liu, Y., and Zhu, H. (2016). ... and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68.

Hirshleifer, D. and Teoh, S. H. (2003). Limited attention, information disclosure, and financial reporting. *Journal of Accounting and Economics*, 36(1-3):337–386.

Hong, H. and Stein, J. C. (1999). A unified theory of underreaction, momentum trading, and overreaction in asset markets. *Journal of Finance*, 54(6):2143–2184.

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.*

Plumlee, M. A. and Maydew, E. L. (2014). The effect of tax-related information complexity on investors' market reactions. *Journal of Accounting Research*, 52(3):563–611.