Cash Earnings Proportion and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

Abstract

This paper studies the asset pricing implications of Cash Earnings Proportion (CEP), 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 CEP achieves an annualized gross (net) Sharpe ratio of 0.45 (0.40), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 21 (16) bps/month with a t-statistic of 2.29 (1.75), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (net income / book equity, Idiosyncratic risk (AHT), Idiosyncratic risk (3 factor), Realized (Total) Volatility, Price, Cash-flow to price variance) is 24 bps/month with a t-statistic of 2.58.

1 Introduction

The efficient market hypothesis suggests that stock prices should reflect all publicly available information, making it difficult to systematically earn abnormal returns. However, a growing body of literature documents persistent patterns in stock returns that appear to contradict market efficiency (Harvey et al., 2016). These patterns, often called anomalies, challenge our understanding of asset pricing and raise important questions about market efficiency.

While numerous return predictors have been documented, many fail to survive transaction costs or more rigorous statistical tests (Hou et al., 2020). This creates an ongoing need to identify robust signals that can predict returns while withstanding careful empirical scrutiny and remaining profitable after accounting for implementation costs.

We propose that a firm's Cash Earnings Proportion (CEP) contains valuable information about future stock returns. The theoretical motivation stems from the quality of earnings literature, which suggests that the cash component of earnings is more persistent and reliable than accruals (Sloan, 1996). When earnings contain a higher proportion of cash-based components, they are likely to be more sustainable and indicative of true economic performance (Dechow et al., 2010).

The relationship between CEP and future returns can be explained through two potential mechanisms. First, firms with higher CEP may be systematically undervalued because investors fail to fully appreciate the superior persistence of cash-based earnings (Richardson et al., 2005). Second, CEP may proxy for firm quality, as companies generating more of their earnings in cash form typically have stronger business models and greater operational efficiency (Novy-Marx, 2013).

These theoretical arguments suggest that stocks with higher CEP should earn superior future returns. This effect should persist even after controlling for known risk factors and related anomalies, as CEP captures a distinct aspect of firm fundamentals that is not fully reflected in current prices (Fama and French, 2015).

Our empirical analysis reveals that CEP strongly predicts cross-sectional stock returns. A value-weighted long-short portfolio that buys stocks with high CEP and shorts those with low CEP generates a monthly alpha of 21 basis points (t-statistic = 2.29) relative to the Fama-French six-factor model. The strategy achieves an annualized gross Sharpe ratio of 0.45, placing it in the top 12% of documented anomalies.

Importantly, the CEP effect remains robust after accounting for transaction costs. The strategy delivers a net Sharpe ratio of 0.40 and maintains statistical significance with a monthly alpha of 16 basis points (t-statistic = 1.75) after incorporating trading costs. This indicates that the anomaly is implementable in practice and not merely a paper trading phenomenon.

The predictive power of CEP persists across different size segments and remains significant among large-cap stocks, with a monthly alpha of 32 basis points (t-statistic = 3.11) in the largest size quintile. This finding is particularly noteworthy as many anomalies fail to generate significant returns among large, liquid stocks where implementation costs are lower.

Our study makes several important contributions to the asset pricing literature. First, we introduce a novel predictor that captures a fundamental aspect of earnings quality not previously explored in cross-sectional return prediction. While prior work has examined various earnings-based signals (Hou et al., 2011), CEP provides unique insights into the cash-generating efficiency of firm operations.

Second, we demonstrate that CEP's predictive power remains robust to controlling for six closely related anomalies, including traditional measures of profitability and risk. The strategy generates a significant alpha of 24 basis points per month (t-statistic = 2.58) even after controlling for these related factors, suggesting it captures distinct information about future returns.

Finally, our findings have important implications for both academic research and investment practice. For academics, we provide new evidence on the role of earnings quality in asset pricing. For practitioners, we identify a robust signal that can enhance portfolio performance, particularly among large-cap stocks where many anomalies fail to generate significant returns.

2 Data

Our study investigates the predictive power of a financial signal derived from accounting data for cross-sectional returns, focusing specifically on the ratio of cash holdings to earnings before interest and taxes (EBIT). 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 CH for cash holdings and item EBIT for earnings. Cash holdings (CH) represent the firm's cash and short-term investments, which are the most liquid assets on a company's balance sheet. EBIT, on the other hand, provides a measure of core operating performance by capturing a firm's profitability before considering financing decisions and tax effects. The construction of the signal follows a straightforward ratio format, where we divide CH by EBIT for each firm in each year of our sample. This ratio, which we term the 'Cash Earnings Proportion,' captures the relative scale of a firm's most liquid assets against its operational income, offering insight into how much of a firm's earnings are held in cash form. By focusing on this relationship, the signal aims to reflect aspects of cash management and operational efficiency in a manner that is both scalable and interpretable. We construct this ratio using end-of-fiscal-year values for both CH and EBIT to ensure consistency and comparability across firms and over time.

3 Signal diagnostics

Figure 1 plots descriptive statistics for the CEP signal. Panel A plots the time-series of the mean, median, and interquartile range for CEP. On average, the cross-sectional mean (median) CEP is 0.82 (0.31) over the 1971 to 2023 sample, where the starting date is determined by the availability of the input CEP data. The signal's interquartile range spans -1.21 to 1.77. Panel B of Figure 1 plots the time-series of the coverage of the CEP signal for the CRSP universe. On average, the CEP signal is available for 6.77% of CRSP names, which on average make up 7.24% of total market capitalization.

4 Does CEP predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on CEP 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 CEP portfolio and sells the low CEP 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 CEP strategy earns an average return of 0.32% per month with a t-statistic of 3.28. The annualized Sharpe ratio of the strategy is 0.45. The alphas range from 0.17% to 0.21% per month and have t-statistics exceeding 1.96 everywhere. The lowest alpha is with respect to the FF3 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 0.18, with a t-statistic of 8.41 on the MKT 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 479 stocks and an average market capitalization of at least $1,177 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 29 bps/month with a t-statistics of 3.34. Out of the twenty-five alphas reported in Panel A, the t-statistics for seventeen exceed two, and for six 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 18-35bps/month. The lowest return, (18 bps/month), is achieved from the quintile sort using NYSE breakpoints and equal-weighted portfolios, and has an associated t-statistic of 1.41. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the CEP trading strategy improves the achievable mean-variance efficient frontier spanned by the factor models in twenty-one cases, and significantly expands the achievable frontier in eight cases.

Table 3 provides direct tests for the role size plays in the CEP strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and CEP, as well as average returns and alphas for long/short trading CEP 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 CEP strategy achieves an average return of 28 bps/month with a t-statistic of 2.56. Among these large cap stocks, the alphas for the CEP strategy relative to the five most common factor models range from 14 to 32 bps/month with t-statistics between 1.36 and 3.11.

5 How does CEP perform relative to the zoo?

Figure 2 puts the performance of CEP 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 CEP strategy falls in the distribution. The CEP strategy's gross (net) Sharpe ratio of 0.45 (0.40) is greater than 88% (97%) of anomaly Sharpe ratios,

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

respectively.

Figure 3 plots the growth of a \$1 invested in these same 212 anomaly trading strategies (gray lines), and compares those with the growth of a \$1 invested in the CEP strategy (red line).[2] Ignoring trading costs, a \$1 invested in the CEP strategy would have yielded \$5.11 which ranks the CEP strategy in the top 3% across the 212 anomalies. Accounting for trading costs, a \$1 invested in the CEP strategy would have yielded \$3.96 which ranks the CEP 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 the CAPM, and the Fama-French three-, four-, five-, and six-factor models from Table 1, and indicates the ranking of the CEP relative to those. Panel A shows that the CEP strategy gross alphas fall between the 38 and 70 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 CEP strategy has a positive net generalized alpha for five out of the five factor models. In these cases CEP ranks between the 61 and 81 percentiles in terms of how much it could have expanded the achievable investment frontier.

[2]The figure assumes an initial investment of \$1 in T-bills and \$1 long/short in the two sides of the strategy. Returns are compounded each month, assuming, as in Detzel et al. (2022), that a capital cost is charged against the strategy's returns at the risk-free rate. This excess return corresponds more closely to the strategy's economic profitability.

6 Does CEP 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 CEP with 210 filtered anomaly signals.[3] Figure 6 also shows an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

A closely related anomaly is also more likely to price CEP or at least to weaken the power CEP has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of CEP conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{CEP} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{CEP} CEP_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{CEP,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 210 filtered anomaly trading strategies at a time. The strategies employed in the spanning tests are constructed using quintile sorts, value-weighting, and NYSE breakpoints. Panel C plots t-statistics on the average returns to strategies constructed by conditional double sorts. In each month, we sort stocks into quintiles based one of the 210 filtered anomaly signals. Then, within each quintile, we sort stocks into quintiles based on CEP. Stocks are finally grouped into five CEP portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted

[3] When performing tests at the underlying signal level (e.g., the correlations plotted in Figure 5), we filter the 212 anomalies to avoid small sample issues. For each anomaly, we calculate the common stock observations in an average month for which both the anomaly and the test signal are available. In the filtered anomaly set, we drop anomalies with fewer than 100 common stock observations in an average month.

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

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on CEP 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 CEP signal in these Fama-MacBeth regressions exceed 0.71, with the minimum t-statistic occurring when controlling for Realized (Total) Volatility. Controlling for all six closely related anomalies, the t-statistic on CEP is 1.16.

Similarly, Table 5 reports results from spanning tests that regress returns to the CEP 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 CEP strategy earns alphas that range from 21-24bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 2.33, which is achieved when controlling for Realized (Total) Volatility. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the CEP trading strategy achieves an alpha of 24bps/month with a t-statistic of 2.58.

7 Does CEP add relative to the whole zoo?

Finally, we can ask how much adding CEP 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 CEP signal.[4] We consider

[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 capital-

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 CEP grows to $1261.16.

8 Conclusion

Our comprehensive analysis of the Cash Earnings Proportion (CEP) signal reveals its significant potential as a predictor of stock returns in the cross-section of equities. The empirical evidence demonstrates that a value-weighted long/short trading strategy based on CEP generates economically meaningful and statistically significant returns, with an impressive annualized gross Sharpe ratio of 0.45 (0.40 net of transaction costs). The strategy's robustness is particularly noteworthy, maintaining significant abnormal returns even after controlling for the Fama-French five factors, momentum, and six closely related anomalies from the factor zoo.

The persistence of the CEP signal's predictive power, evidenced by a monthly alpha of 24 basis points (t-statistic of 2.58) after controlling for multiple factors, suggests that it captures unique information about future stock returns not explained by existing factors. This finding has important implications for both academic research and practical investment management, as it contributes to our understanding of market efficiency and offers potential opportunities for portfolio enhancement.

However, several limitations should be considered. First, our analysis focuses on a specific time period, and the signal's effectiveness may vary across different market

ization on CRSP in the period for which CEP is available.

conditions. Second, transaction costs and market impact could affect the strategy's real-world implementation, particularly for smaller investors or larger portfolios.

Future research could explore the international validity of the CEP signal, its interaction with other market anomalies, and its performance during different economic cycles. Additionally, investigating the underlying economic mechanisms driving the CEP premium would provide valuable insights into market behavior and asset pricing theory. Researchers might also consider examining how the signal's effectiveness varies across different market capitalizations and sectors.

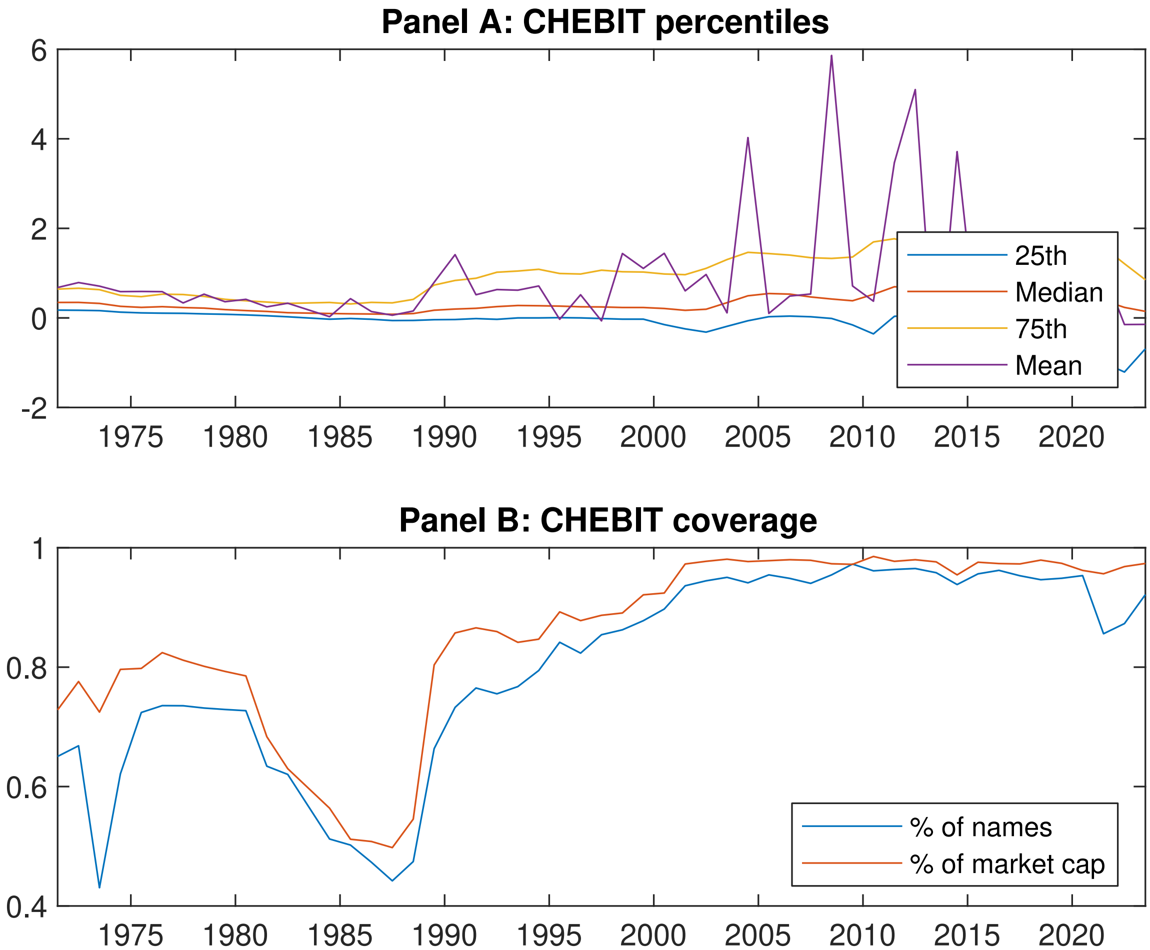

Figure 1: Times series of CEP percentiles and coverage.
This figure plots descriptive statistics for CEP. Panel A shows cross-sectional percentiles of CEP over the sample. Panel B plots the monthly coverage of CEP relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on CEP. 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 CEP-sorted portfolios						
	(L)	(2)	(3)	(4)	(H)	(H-L)
r^e	0.45	0.57	0.62	0.66	0.77	0.32
	[2.33]	[3.33]	[3.39]	[3.28]	[3.34]	[3.28]
α_{CAPM}	-0.15	0.04	0.04	0.02	0.04	0.19
	[-2.25]	[0.69]	[0.87]	[0.45]	[0.68]	[2.12]
α_{FF3}	-0.13	0.02	0.02	0.02	0.04	0.17
	[-2.02]	[0.47]	[0.38]	[0.39]	[0.81]	[1.96]
α_{FF4}	-0.13	0.02	0.01	0.05	0.06	0.19
	[-2.06]	[0.39]	[0.24]	[0.99]	[1.00]	[2.10]
α_{FF5}	-0.05	-0.13	-0.09	-0.00	0.16	0.20
	[-0.73]	[-2.84]	[-1.95]	[-0.01]	[2.90]	[2.21]
α_{FF6}	-0.05	-0.12	-0.08	0.03	0.16	0.21
	[-0.85]	[-2.62]	[-1.87]	[0.50]	[2.89]	[2.29]
Panel B: Fama and French (2018) 6-factor model loadings for CEP-sorted portfolios						
β_{MKT}	0.95	0.96	1.01	1.05	1.13	0.18
	[64.73]	[91.21]	[96.99]	[86.38]	[89.74]	[8.41]
β_{SMB}	0.06	-0.09	0.01	0.05	0.19	0.13
	[2.65]	[-5.76]	[0.76]	[2.53]	[10.21]	[4.20]
β_{HML}	-0.05	-0.02	0.01	-0.02	-0.00	0.05
	[-1.71]	[-1.13]	[0.54]	[-0.67]	[-0.03]	[1.16]
β_{RMW}	-0.23	0.31	0.22	0.08	-0.24	-0.01
	[-7.91]	[15.36]	[11.00]	[3.22]	[-9.70]	[-0.28]
β_{CMA}	-0.00	0.16	0.10	-0.00	-0.08	-0.08
	[-0.04]	[5.11]	[3.38]	[-0.07]	[-2.29]	[-1.32]
β_{UMD}	0.01	-0.01	-0.00	-0.04	-0.00	-0.02
	[0.88]	[-1.20]	[-0.35]	[-3.31]	[-0.19]	[-0.72]
Panel C: Average number of firms (n) and market capitalization (me)						
n	1182	479	533	624	948	
me ($\$10^6$)	1177	2488	2891	2232	1783	

Table 2: Robustness to sorting methodology & trading costs

This table evaluates the robustness of the choices made in the CEP 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	0.32	0.19	0.17	0.19	0.20	0.21
			[3.28]	[2.12]	[1.96]	[2.10]	[2.21]	[2.29]
Quintile	NYSE	EW	0.39	0.46	0.36	0.24	0.10	0.01
			[3.13]	[3.70]	[3.15]	[2.08]	[1.01]	[0.12]
Quintile	Name	VW	0.39	0.47	0.32	0.27	-0.01	-0.04
			[2.52]	[3.02]	[2.27]	[1.83]	[-0.11]	[-0.34]
Quintile	Cap	VW	0.29	0.17	0.17	0.19	0.28	0.28
			[3.34]	[2.12]	[2.25]	[2.48]	[3.55]	[3.62]
Decile	NYSE	VW	0.41	0.39	0.31	0.31	0.15	0.16
			[3.25]	[3.13]	[2.56]	[2.50]	[1.30]	[1.33]
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	0.28	0.16	0.14	0.15	0.15	0.16
			[2.93]	[1.74]	[1.57]	[1.69]	[1.62]	[1.75]
Quintile	NYSE	EW	0.18	0.23	0.13	0.06		
			[1.41]	[1.76]	[1.10]	[0.53]		
Quintile	Name	VW	0.34	0.42	0.29	0.26		
			[2.17]	[2.68]	[2.02]	[1.78]		
Quintile	Cap	VW	0.26	0.14	0.14	0.16	0.21	0.22
			[3.00]	[1.73]	[1.82]	[2.00]	[2.75]	[2.88]
Decile	NYSE	VW	0.35	0.34	0.26	0.26	0.11	0.13
			[2.82]	[2.71]	[2.18]	[2.17]	[0.96]	[1.11]

Table 3: Conditional sort on size and CEP

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

		CEP Quintiles					CEP Strategies					
		(L)	(2)	(3)	(4)	(H)	r^e	α_{CAPM}	α_{FF3}	α_{FF4}	α_{FF5}	α_{FF6}
Size quintiles	(1)	0.32	0.28	0.70	0.82	0.90	0.58	0.71	0.56	0.49	0.21	0.17
		[0.99]	[0.87]	[2.79]	[3.39]	[3.45]	[3.65]	[4.58]	[4.02]	[3.50]	[1.69]	[1.42]
	(2)	0.45	0.68	0.80	0.83	0.77	0.32	0.42	0.28	0.25	0.02	0.00
		[1.47]	[2.84]	[3.32]	[3.44]	[3.02]	[2.39]	[3.19]	[2.37]	[2.04]	[0.20]	[0.03]
	(3)	0.53	0.64	0.76	0.86	0.77	0.23	0.26	0.16	0.02	-0.08	-0.18
		[2.01]	[2.92]	[3.33]	[3.67]	[3.15]	[1.90]	[2.14]	[1.35]	[0.18]	[-0.71]	[-1.58]
	(4)	0.54	0.65	0.71	0.72	0.88	0.34	0.26	0.24	0.18	0.21	0.17
		[2.44]	[3.22]	[3.35]	[3.22]	[3.56]	[3.25]	[2.51]	[2.32]	[1.76]	[2.02]	[1.60]
	(5)	0.42	0.56	0.58	0.54	0.71	0.28	0.14	0.16	0.20	0.30	0.32
		[2.37]	[3.34]	[3.19]	[2.72]	[3.25]	[2.56]	[1.36]	[1.55]	[1.89]	[2.95]	[3.11]

Panel B: Portfolio average number of firms and market capitalization

		CEP Quintiles					CEP Quintiles				
		Average n					Average market capitalization ($\$10^6$)				
		(L)	(2)	(3)	(4)	(H)	(L)	(2)	(3)	(4)	(H)
Size quintiles	(1)	435	424	441	444	444	34	29	42	45	42
	(2)	115	115	116	116	115	62	63	66	66	65
	(3)	78	79	79	79	79	106	109	111	108	107
	(4)	64	64	64	64	64	218	229	228	228	229
	(5)	57	58	58	58	57	1200	2006	2005	1638	1536

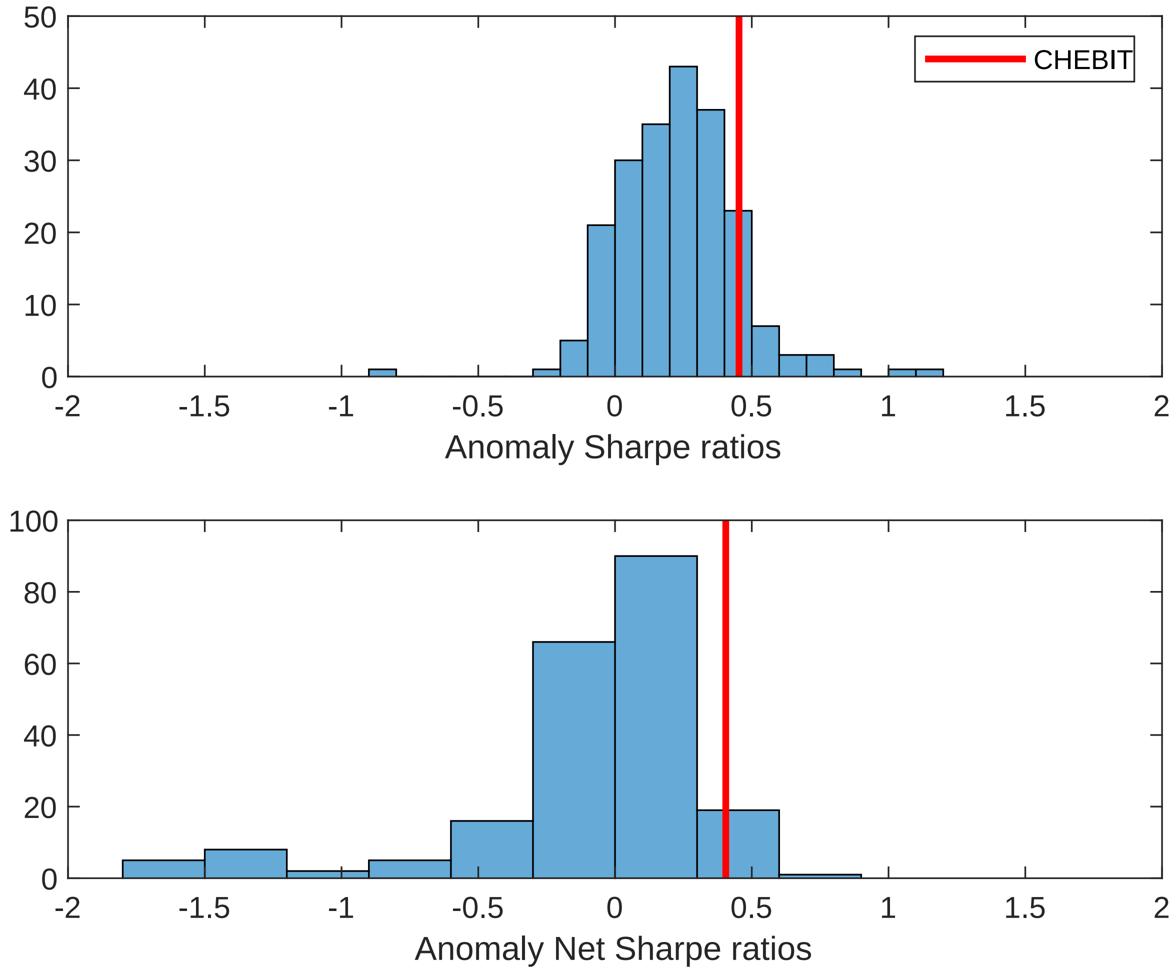

Figure 2: Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the CEP with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

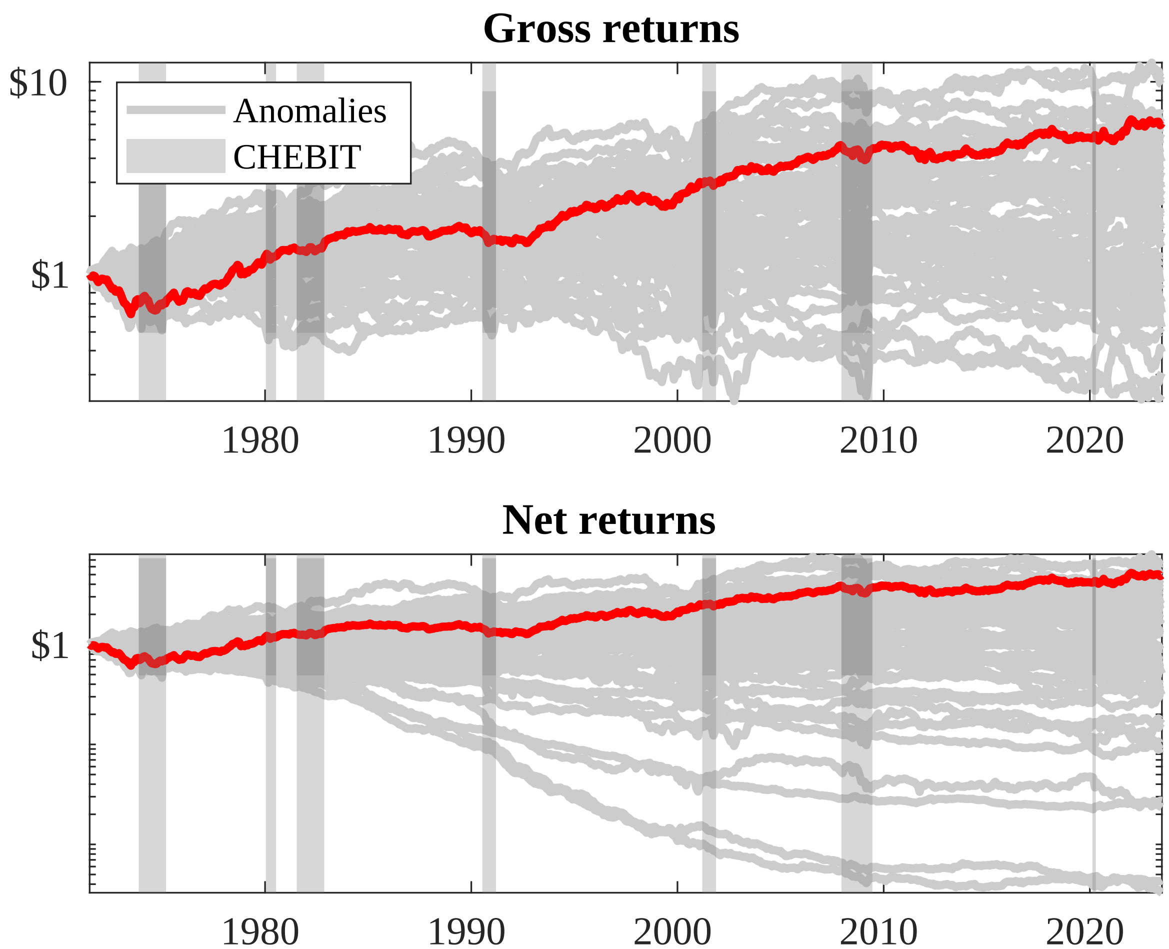

Figure 3: Dollar invested.
This figure plots the growth of a $1 invested in 212 anomaly trading strategies (gray lines), and compares those with the CEP trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for gross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

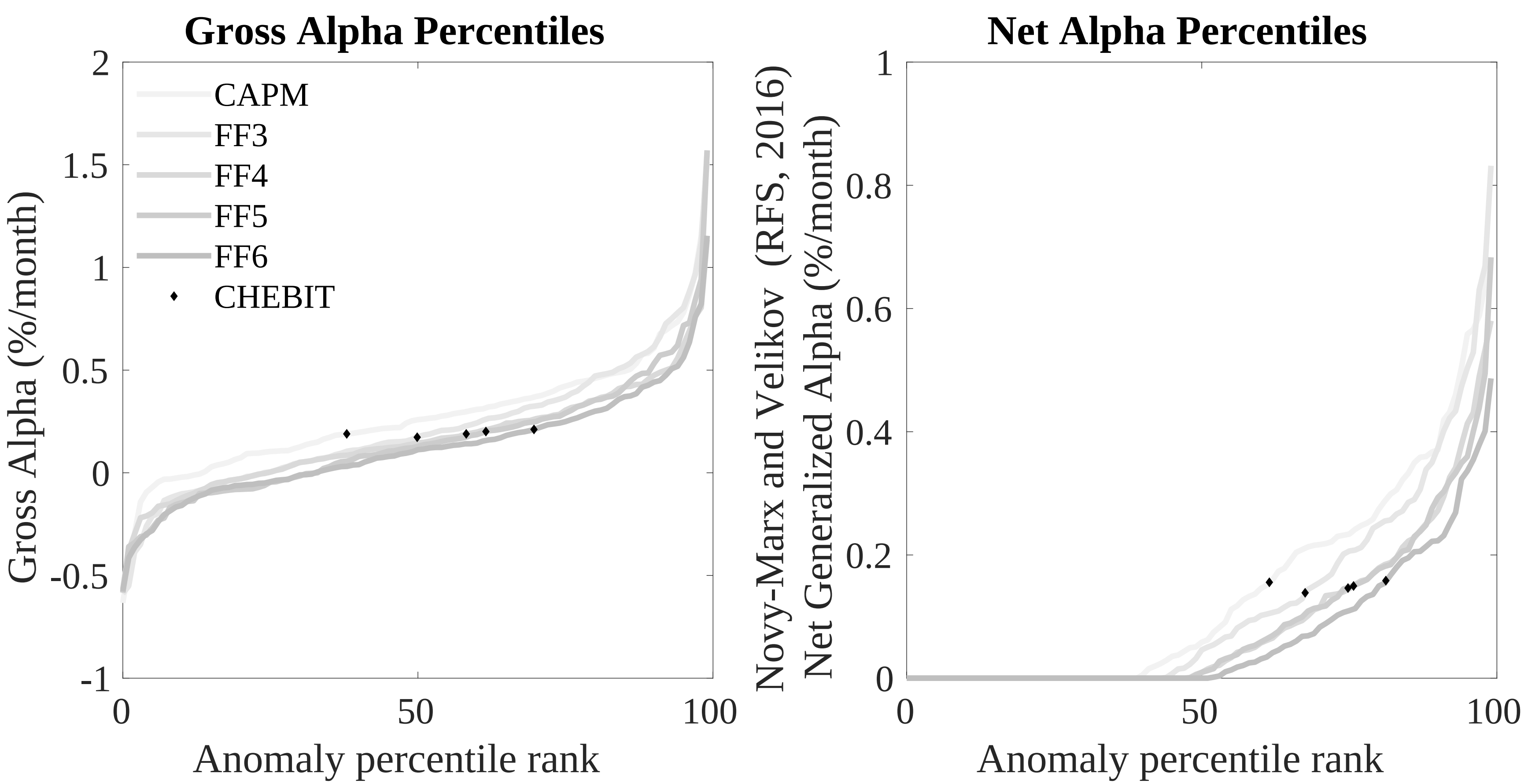

Figure 4: Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the CEP trading strategy alphas (diamonds). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. The alphas include those with respect to the CAPM, Fama and French (1993) three-factor model, Fama and French (1993) three-factor model augmented with the Carhart (1997) momentum factor, Fama and French (2015) five-factor model, and the Fama and French (2015) five-factor model augmented with the Carhart (1997) momentum factor following Fama and French (2018). The left panel plots alphas with no adjustment for trading costs. The right panel plots Novy-Marx and Velikov (2016) net generalized alphas.

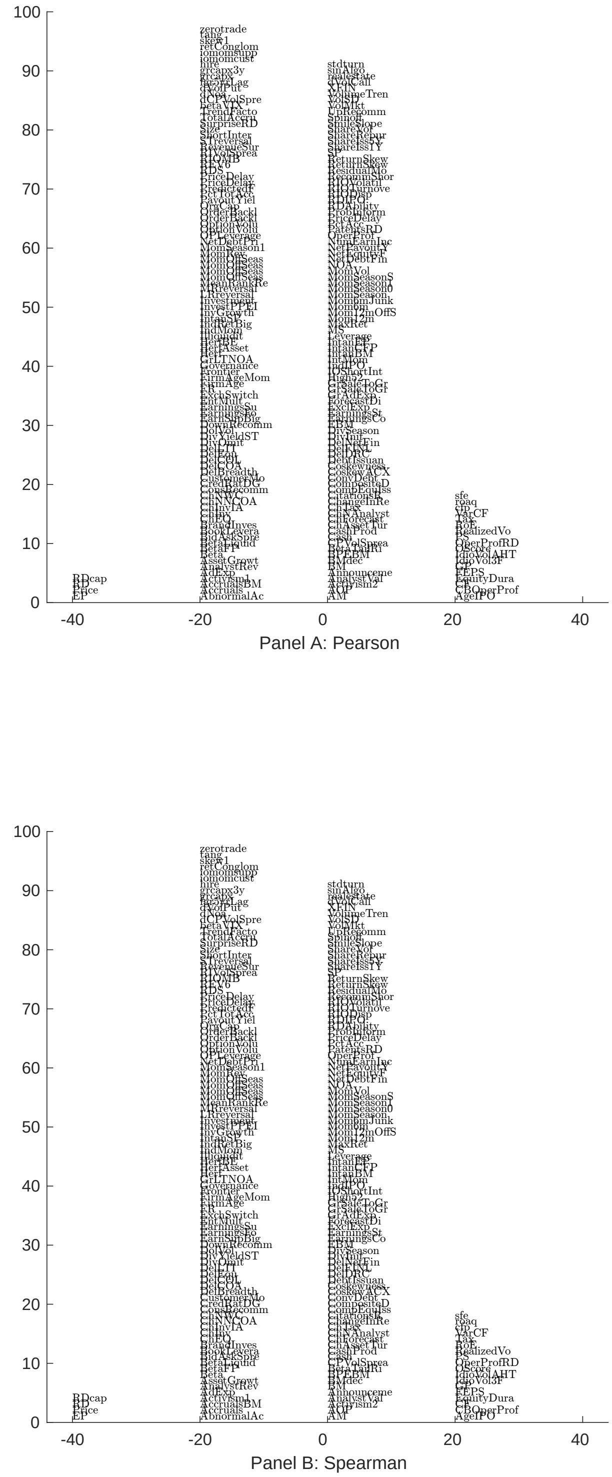

Figure 5: Distribution of correlations.
This figure plots a name histogram of correlations of 210 filtered anomaly signals with CEP. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

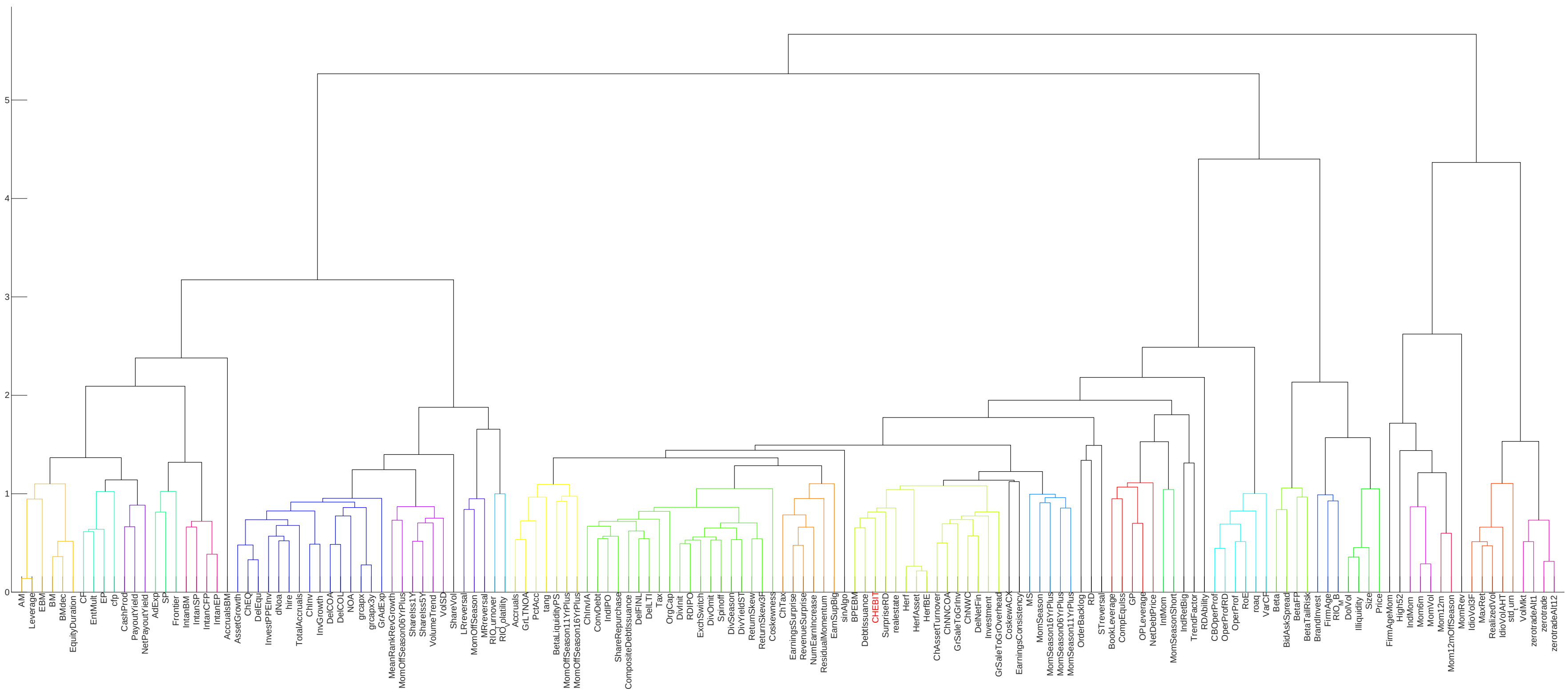

Figure 6: Agglomerative hierarchical cluster plot

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

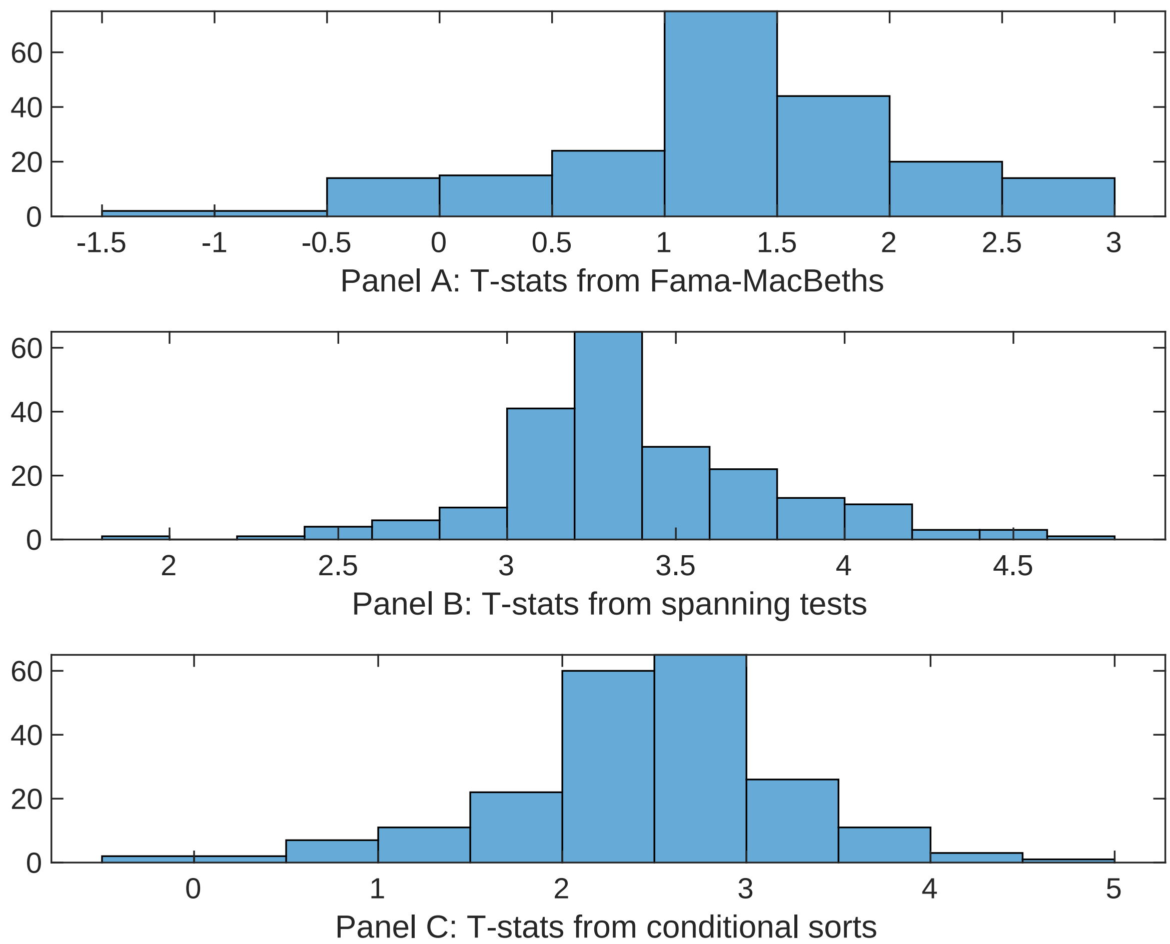

Figure 7: Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of CEP conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on β_{CEP} from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{CEP}CEP_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where X stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on α from spanning tests of the form: $r_{CEP,t} = \alpha + \beta r_{X,t} + \epsilon_t$, where $r_{X,t}$ stands for the returns to one of the 210 filtered anomaly trading strategies at a time. The strategies employed in the spanning tests are constructed using quintile sorts, value-weighting, and NYSE breakpoints. Panel C plots t-statistics on the average returns to strategies constructed by conditional double sorts. In each month, we sort stocks into quintiles based one of the 210 filtered anomaly signals at a time. Then, within each quintile, we sort stocks into quintiles based on CEP. Stocks are finally grouped into five CEP portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted CEP trading strategies conditioned on each of the 210 filtered anomalies.

Table 4: Fama-MacBeths controlling for most closely related anomalies
This table presents Fama-MacBeth results of returns on CEP. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{CEP} CEP_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^{k} + \epsilon_{i,t}$. The six most closely related anomalies, X, are net income / book equity, Idiosyncratic risk (AHT), Idiosyncratic risk (3 factor), Realized (Total) Volatility, Price, Cash-flow to price variance. 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	0.12	0.13	0.14	0.15	0.12	0.12	0.17
	[4.67]	[6.93]	[7.50]	[8.16]	[4.06]	[5.19]	[8.38]
CEP	0.72	0.51	0.49	0.43	0.70	0.14	0.69
	[1.11]	[0.86]	[0.81]	[0.71]	[1.10]	[1.91]	[1.16]
Anomaly 1	0.49						0.96
	[0.04]						[0.01]
Anomaly 2		0.89					-0.21
		[1.62]					[-0.49]
Anomaly 3			0.15				-0.15
			[3.48]				[-1.32]
Anomaly 4				0.14			0.31
				[3.46]			[2.75]
Anomaly 5					0.58		0.87
					[1.20]		[2.99]
Anomaly 6						-0.47	-0.90
						[-0.57]	[-2.69]
# months	624	619	619	619	624	619	619
$\bar{R}^2$(%)	0	2	1	2	1	1	0

Table 5: Spanning tests controlling for most closely related anomalies
This table presents spanning tests results of regressing returns to the CEP trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{CEP} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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 income / book equity, Idiosyncratic risk (AHT), Idiosyncratic risk (3 factor), Realized (Total) Volatility, Price, Cash-flow to price variance. 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	0.22	0.23	0.24	0.23	0.21	0.23	0.24
	[2.39]	[2.50]	[2.58]	[2.47]	[2.33]	[2.46]	[2.58]
Anomaly 1	-17.23						-15.08
	[-3.39]						[-2.68]
Anomaly 2		-6.31					-0.20
		[-2.04]					[-0.04]
Anomaly 3			-7.97				-0.03
			[-2.51]				[-0.01]
Anomaly 4				-7.57			-7.25
				[-2.85]			[-1.48]
Anomaly 5					-1.86		-6.16
					[-0.50]		[-1.44]
Anomaly 6						-4.05	1.67
						[-1.14]	[0.38]
mkt	15.99	15.83	15.53	14.62	18.98	17.10	13.88
	[7.07]	[6.45]	[6.48]	[5.88]	[8.54]	[7.18]	[5.40]
smb	8.35	8.05	7.49	8.81	16.14	11.23	11.50
	[2.31]	[1.83]	[1.79]	[2.37]	[3.18]	[2.71]	[2.15]
hml	3.57	7.87	8.29	8.52	5.86	4.21	8.28
	[0.88]	[1.88]	[1.98]	[2.04]	[1.43]	[0.96]	[1.68]
rmw	13.26	3.00	2.90	3.27	-3.42	-1.49	12.97
	[2.14]	[0.59]	[0.61]	[0.70]	[-0.74]	[-0.34]	[1.99]
cma	-11.43	-7.57	-7.03	-6.33	-7.76	-10.84	-8.57
	[-1.84]	[-1.20]	[-1.11]	[-1.00]	[-1.25]	[-1.71]	[-1.27]
umd	-1.43	-0.53	-0.15	-0.02	-2.69	-1.02	-3.63
	[-0.68]	[-0.24]	[-0.07]	[-0.01]	[-0.90]	[-0.47]	[-1.20]
# months	624	620	620	620	624	620	620
$\bar{R}^2$(%)	22	21	21	21	20	20	22

Panel A: Average rank

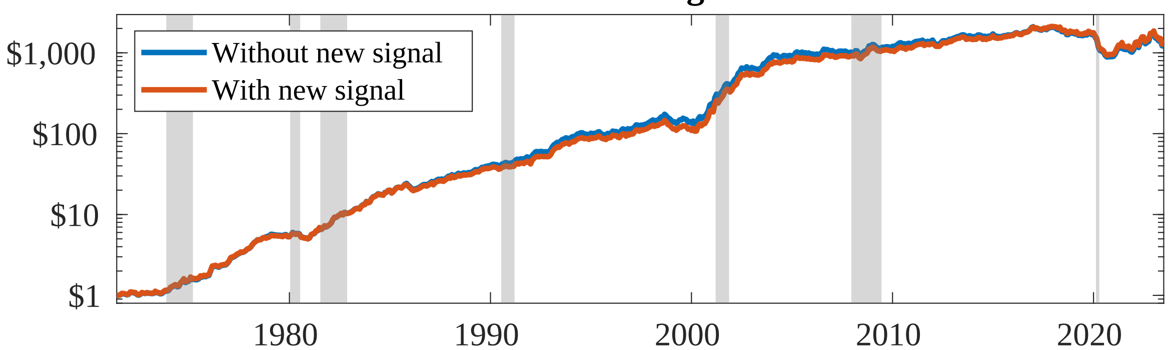

Figure 8: Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 156 anomalies. The red solid lines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

Dechow, P., Ge, W., and Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, 50(2-3):344–401.

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.

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

Hou, K., Xue, C., and Zhang, L. (2011). Dissecting the effect of the cross-section of expected stock returns. *Review of Financial Studies*, 24(4):1152–1191.

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

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

Novy-Marx, R. and Velikov, M. (2016). A taxonomy of anomalies and their trading costs. *Review of Financial Studies*, 29(1):104–147.

Novy-Marx, R. and Velikov, M. (2023). Assaying anomalies. *Working paper.*

Richardson, S. A., Sloan, R. G., Soliman, M. T., and Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics*, 39(3):437–485.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3):289–315.