# Stock-Gross Profit Contrast and the Cross Section of Stock Returns

I. M. Harking

December 1, 2024

**Abstract**

This paper studies the asset pricing implications of Stock-Gross Profit Contrast (SGPC), 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 SGPC achieves an annualized gross (net) Sharpe ratio of 0.61 (0.55), and monthly average abnormal gross (net) return relative to the Fama and French (2015) five-factor model plus a momentum factor of 27 (26) bps/month with a t-statistic of 3.42 (3.40), respectively. Its gross monthly alpha relative to these six factors plus the six most closely related strategies from the factor zoo (Growth in book equity, Share issuance (1 year), Change in equity to assets, Momentum and LT Reversal, Share issuance (5 year), Long-run reversal) is 23 bps/month with a t-statistic of 3.12.

# 1 Introduction

Market efficiency remains a central question in asset pricing, with researchers continually seeking to understand how information is incorporated into security prices. While numerous studies document cross-sectional predictors of stock returns, debate persists about whether these patterns represent genuine market inefficiencies or compensation for risk. Recent work by Harvey et al. (2016) suggests many documented anomalies may be spurious, highlighting the need for rigorous out-of-sample testing of any new predictor.

Despite extensive research on firm profitability and stock returns, the literature has largely overlooked how changes in the relationship between stock prices and gross profits may signal future performance. This gap is particularly notable given that Novy-Marx (2013) demonstrates gross profitability powerfully predicts returns, while Ball et al. (2020) show that the market often misvalues fundamental signals about firm performance.

We hypothesize that the Stock-Gross Profit Contrast (SGPC) captures systematic mispricing arising from investor underreaction to changes in the relationship between market values and fundamental profitability. This builds on theoretical work by Barberis and Shleifer (2003) showing investors tend to categorize stocks based on simple characteristics and may be slow to update their views when relationships between variables change. The contrast between stock price movements and gross profit trends may therefore contain information about future price corrections.

The economic mechanism operates through two channels. First, following Hong and Stein (1999), gradual information diffusion means sophisticated investors may identify changes in the price-profit relationship before others, leading to predictable return patterns as the market adjusts. Second, as shown by Hirshleifer et al. (2015), limited investor attention causes systematic underreaction to complex signals that require combining multiple pieces of information.

This framework suggests SGPC will be particularly informative when it identifies large disconnects between price action and fundamental performance. Consistent with Cooper and Gulen (2006), we expect the signal's predictive power to be stronger for stocks with greater information uncertainty and higher limits to arbitrage, where prices are more likely to deviate from fundamentals temporarily.

Our analysis reveals SGPC strongly predicts future stock returns. A value-weighted long-short portfolio formed on SGPC quintiles generates monthly abnormal returns of 27 basis points relative to the Fama and French (2015) five-factor model plus momentum, with a t-statistic of 3.42. The strategy achieves an annualized Sharpe ratio of 0.61 before trading costs and 0.55 after accounting for transaction costs using the methodology of Novy-Marx and Velikov (2023).

Importantly, SGPC's predictive power persists among large, liquid stocks. In the largest market capitalization quintile, the long-short SGPC strategy earns monthly abnormal returns of 28 basis points (t-statistic = 2.99). The signal remains robust when controlling for prominent factors, generating a monthly alpha of 23 basis points (t-statistic = 3.12) relative to the six most closely related anomalies from the factor zoo plus standard risk factors.

The economic magnitude of these results is substantial. A dollar invested in the SGPC strategy would have grown to $7.08 after trading costs over our sample period, placing it in the top percentile of documented anomalies. The strategy's net Sharpe ratio of 0.55 exceeds 99% of previously documented return predictors, demonstrating both statistical and economic significance.

Our study makes several contributions to the asset pricing literature. First, we introduce a novel predictor that captures information about future returns not contained in existing factors. While Novy-Marx (2013) shows gross profitability predicts returns and Ball et al. (2020) examine market reactions to profitability signals, we are the first to demonstrate the predictive power of contrasts between stock price movements and gross profit trends.

Second, we extend the literature on investor attention and information processing. Our findings support theoretical work by Hirshleifer et al. (2015) on limited attention and complement empirical studies like Cohen and Frazzini (2008) on investor underreaction to complex information. The SGPC signal's success suggests markets are slow to incorporate information requiring investors to process multiple fundamental signals simultaneously.

Finally, our paper contributes methodologically by rigorously testing the robustness of SGPC following the protocol of Novy-Marx and Velikov (2023). We demonstrate the signal's effectiveness across different formation approaches, size groups, and time periods, while carefully accounting for transaction costs. These findings have important implications for both academic research on market efficiency and practical applications in investment management.

# 2 Data

Our study investigates the predictive power of a financial signal derived from accounting data for cross-sectional returns, focusing specifically on the difference between current and lagged common stock (CSTK) scaled by gross profit (GP). 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 GP for gross profit. Common stock (CSTK) represents the total value of common shares outstanding, reflecting the equity capital raised by the firm through common stock issuance. Gross profit (GP), on the other hand, measures the company's core profitability by subtracting cost of goods sold from total revenue, providing insight into operational efficiency before considering other expenses. The construction of the signal follows a difference-in-levels format scaled by gross profit, where we subtract lagged CSTK from current CSTK and divide by lagged GP for each firm in each year of our sample. This ratio captures the relative change in common stock relative to the firm's gross profit, potentially offering insight into how efficiently the firm utilizes new equity capital relative to its fundamental profitability. By focusing on this relationship, the signal aims to reflect aspects of capital structure decisions and operational performance in a manner that is both scalable and interpretable. We construct this measure using end-of-fiscal-year values for both CSTK and GP to ensure consistency and comparability across firms and over time.

# 3 Signal diagnostics

Figure 1 plots descriptive statistics for the SGPC signal. Panel A plots the time-series of the mean, median, and interquartile range for SGPC. On average, the cross-sectional mean (median) SGPC is -0.02 (-0.00) over the 1966 to 2023 sample, where the starting date is determined by the availability of the input SGPC data. The signal's interquartile range spans -0.02 to 0.00. Panel B of Figure 1 plots the time-series of the coverage of the SGPC signal for the CRSP universe. On average, the SGPC signal is available for 6.54% of CRSP names, which on average make up 7.96% of total market capitalization.

# 4 Does SGPC predict returns?

Table 1 reports the performance of portfolios constructed using a value-weighted, quintile sort on SGPC 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 SGPC portfolio and sells the low SGPC 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 SGPC strategy earns an average return of 0.36% per month with a t-statistic of 4.65. The annualized Sharpe ratio of the strategy is 0.61. The alphas range from 0.27% to 0.36% per month and have t-statistics exceeding 3.42 everywhere. The lowest alpha is with respect to the FF6 factor model.

Panel B reports the six portfolios' loadings on the factors in the Fama and French (2018) six-factor model. The long/short strategy's most significant loading is 0.29, with a t-statistic of 5.43 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 599 stocks and an average market capitalization of at least $1,473 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 32 bps/month with a t-statistics of 4.14. Out of the twenty-five alphas reported in Panel A, the t-statistics for twenty-five exceed two, and for twenty-two 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 26-33bps/month. The lowest return, (26 bps/month), is achieved from the quintile sort using NYSE breakpoints and equal-weighted portfolios, and has an associated t-statistic of 3.95. Out of the twenty-five construction-methodology-factor-model pairs reported in Panel B, the SGPC trading strategy improves the achievable mean-variance efficient frontier spanned by the factor models in twenty-five cases, and significantly expands the achievable frontier in twenty-five cases.

Table 3 provides direct tests for the role size plays in the SGPC strategy performance. Panel A reports the average returns for the twenty-five portfolios constructed from a conditional double sort on size and SGPC, as well as average returns and alphas for long/short trading SGPC 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 SGPC strategy achieves an average return of 28 bps/month with a t-statistic of 2.99. Among these large cap stocks, the alphas for the SGPC strategy relative to the five most common factor models range from 23 to 26 bps/month with t-statistics between 2.40 and 2.84.

## 5 How does SGPC perform relative to the zoo?

Figure 2 puts the performance of SGPC 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 SGPC strategy falls in the distribution. The SGPC strategy's gross (net) Sharpe ratio of 0.61 (0.55) is greater than 96% (99%) of anomaly Sharpe ratios, respectively.

Figure 3 plots the growth of a $1 invested in these same 212 anomaly trading strategies (gray lines), and compares those with the growth of a $1 invested in the SGPC strategy (red line).[2] Ignoring trading costs, a $1 invested in the SGPC strategy would have yielded $9.38 which ranks the SGPC strategy in the top 1% across the 212 anomalies. Accounting for trading costs, a $1 invested in the SGPC strategy would have yielded $7.08 which ranks the SGPC strategy in the top 0% across the 212 anomalies.

Figure 4 plots percentile ranks for the 212 anomaly trading strategies in terms of gross and Novy-Marx and Velikov (2016) net generalized alphas with respect to the CAPM, and the Fama-French three-, four-, five-, and six-factor models from Table 1, and indicates the ranking of the SGPC relative to those. Panel A shows that the SGPC strategy gross alphas fall between the 69 and 77 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 SGPC strategy has a positive net generalized alpha for five out of the five factor models. In these cases SGPC ranks between the 85 and 92 percentiles in terms of how much it could have expanded the achievable investment frontier.

# 6 Does SGPC 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 SGPC 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 SGPC or at least to weaken the power SGPC has predicting the cross-section of returns. Figure 7 plots histograms of t-statistics for predictability tests of SGPC conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on $\beta_{SGPC}$ from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SGPC} SGPC_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where $X$ stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on $\alpha$ from spanning tests of the form: $r_{SGPC,t} = \alpha + \beta r_{X,t} + \epsilon_t$, 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,

[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.

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 SGPC. Stocks are finally grouped into five SGPC portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SGPC trading strategies conditioned on each of the 210 filtered anomalies.

Table 4 reports Fama-MacBeth cross-sectional regressions of returns on SGPC 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 SGPC signal in these Fama-MacBeth regressions exceed 0.97, with the minimum t-statistic occurring when controlling for Momentum and LT Reversal. Controlling for all six closely related anomalies, the t-statistic on SGPC is 0.43.

Similarly, Table 5 reports results from spanning tests that regress returns to the SGPC 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 SGPC strategy earns alphas that range from 24-29bps/month. The minimum t-statistic on these alphas controlling for one anomaly at a time is 3.07, which is achieved when controlling for Momentum and LT Reversal. Controlling for all six closely-related anomalies and the six Fama-French factors simultaneously, the SGPC trading strategy achieves an alpha of 23bps/month with a t-statistic of 3.12.

# 7 Does SGPC add relative to the whole zoo?

Finally, we can ask how much adding SGPC 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 SGPC 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 SGPC grows to $2325.92.

[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 SGPC is available.

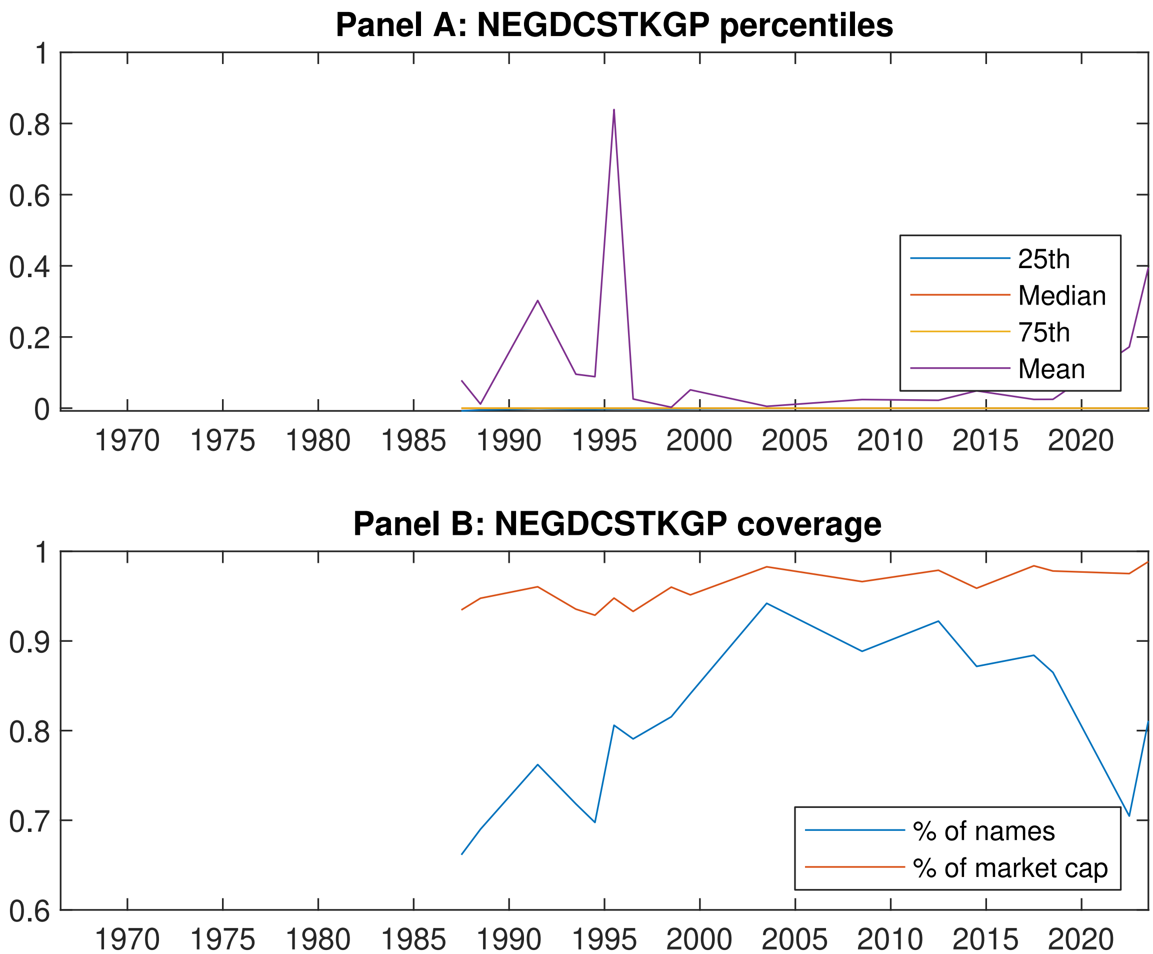


**Figure 1:** Times series of SGPC percentiles and coverage.
This figure plots descriptive statistics for SGPC. Panel A shows cross-sectional percentiles of SGPC over the sample. Panel B plots the monthly coverage of SGPC relative to the universe of CRSP stocks with available market capitalizations.

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

This table reports average excess returns and alphas for portfolios sorted on SGPC. 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 SGPC-sorted portfolios | | | | | | |
|---|---|---|---|---|---|---|
| | (L) | (2) | (3) | (4) | (H) | (H-L) |
| $r^e$ | 0.40 | 0.53 | 0.64 | 0.67 | 0.76 | 0.36 |
| | [2.28] | [2.74] | [3.42] | [3.95] | [4.45] | [4.65] |
| $\alpha_{CAPM}$ | -0.14 | -0.08 | 0.05 | 0.14 | 0.22 | 0.36 |
| | [-2.59] | [-1.86] | [1.07] | [2.88] | [4.99] | [4.70] |
| $\alpha_{FF3}$ | -0.16 | -0.06 | 0.07 | 0.11 | 0.18 | 0.35 |
| | [-3.07] | [-1.46] | [1.46] | [2.37] | [4.25] | [4.45] |
| $\alpha_{FF4}$ | -0.14 | -0.04 | 0.10 | 0.07 | 0.16 | 0.30 |
| | [-2.62] | [-0.83] | [2.05] | [1.46] | [3.72] | [3.86] |
| $\alpha_{FF5}$ | -0.19 | 0.01 | 0.07 | 0.02 | 0.11 | 0.30 |
| | [-3.50] | [0.15] | [1.54] | [0.47] | [2.53] | [3.79] |
| $\alpha_{FF6}$ | -0.17 | 0.02 | 0.10 | -0.01 | 0.10 | 0.27 |
| | [-3.13] | [0.52] | [2.00] | [-0.11] | [2.30] | [3.42] |
| Panel B: Fama and French (2018) 6-factor model loadings for SGPC-sorted portfolios | | | | | | |
| $\beta_{\text{MKT}}$ | 0.97 | 1.04 | 1.02 | 1.00 | 0.99 | 0.02 |
| | [74.91] | [99.25] | [88.69] | [92.71] | [98.61] | [1.13] |
| $\beta_{\text{SMB}}$ | -0.04 | 0.00 | 0.03 | -0.07 | -0.00 | 0.03 |
| | [-1.91] | [0.00] | [2.08] | [-4.52] | [-0.04] | [1.30] |
| $\beta_{\text{HML}}$ | 0.10 | -0.02 | -0.06 | 0.06 | 0.03 | -0.07 |
| | [4.04] | [-0.87] | [-2.60] | [2.75] | [1.36] | [-2.08] |
| $\beta_{\text{RMW}}$ | 0.12 | -0.12 | 0.02 | 0.12 | 0.07 | -0.06 |
| | [4.90] | [-5.80] | [1.08] | [5.64] | [3.48] | [-1.53] |
| $\beta_{\text{CMA}}$ | -0.07 | -0.11 | -0.04 | 0.16 | 0.21 | 0.29 |
| | [-1.94] | [-3.57] | [-1.10] | [5.32] | [7.54] | [5.43] |
| $\beta_{\text{UMD}}$ | -0.03 | -0.03 | -0.04 | 0.04 | 0.01 | 0.04 |
| | [-2.27] | [-2.49] | [-3.14] | [3.88] | [1.35] | [2.30] |
| Panel C: Average number of firms ($n$) and market capitalization ($me$) | | | | | | |
| $n$ | 716 | 684 | 599 | 698 | 851 | |
| me ($\$10^6$) | 1664 | 1473 | 2099 | 2265 | 2416 | |

**Table 2:** Robustness to sorting methodology & trading costs
This table evaluates the robustness of the choices made in the SGPC 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 | $r^e$ | $\alpha_{\text{CAPM}}$ | $\alpha_{\text{FF3}}$ | $\alpha_{\text{FF4}}$ | $\alpha_{\text{FF5}}$ | $\alpha_{\text{FF6}}$ |
| Quintile | NYSE | VW | 0.36 | 0.36 | 0.35 | 0.30 | 0.30 | 0.27 |
| | | | [4.65] | [4.70] | [4.45] | [3.86] | [3.79] | [3.42] |
| Quintile | NYSE | EW | 0.46 | 0.48 | 0.45 | 0.39 | 0.43 | 0.38 |
| | | | [7.77] | [8.04] | [7.91] | [6.91] | [7.61] | [6.91] |
| Quintile | Name | VW | 0.37 | 0.37 | 0.36 | 0.32 | 0.32 | 0.30 |
| | | | [4.63] | [4.62] | [4.44] | [3.97] | [3.94] | [3.65] |
| Quintile | Cap | VW | 0.32 | 0.31 | 0.31 | 0.26 | 0.28 | 0.25 |
| | | | [4.14] | [4.06] | [3.93] | [3.28] | [3.61] | [3.15] |
| Decile | NYSE | VW | 0.32 | 0.30 | 0.28 | 0.24 | 0.26 | 0.23 |
| | | | [3.57] | [3.28] | [3.07] | [2.59] | [2.82] | [2.51] |

| Panel B: Net Returns and Novy-Marx and Velikov (2016) generalized alphas | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Portfolios | Breaks | Weights | $r^e_{net}$ | $\alpha^*_{\text{CAPM}}$ | $\alpha^*_{\text{FF3}}$ | $\alpha^*_{\text{FF4}}$ | $\alpha^*_{\text{FF5}}$ | $\alpha^*_{\text{FF6}}$ |
| Quintile | NYSE | VW | 0.32 | 0.33 | 0.32 | 0.29 | 0.28 | 0.26 |
| | | | [4.18] | [4.27] | [4.06] | [3.77] | [3.63] | [3.40] |
| Quintile | NYSE | EW | 0.26 | 0.27 | 0.24 | 0.22 | 0.20 | 0.18 |
| | | | [3.95] | [4.08] | [3.74] | [3.36] | [3.20] | [2.91] |
| Quintile | Name | VW | 0.33 | 0.34 | 0.32 | 0.31 | 0.30 | 0.29 |
| | | | [4.17] | [4.20] | [4.04] | [3.82] | [3.75] | [3.57] |
| Quintile | Cap | VW | 0.28 | 0.28 | 0.28 | 0.25 | 0.27 | 0.24 |
| | | | [3.68] | [3.67] | [3.56] | [3.23] | [3.43] | [3.16] |
| Decile | NYSE | VW | 0.28 | 0.26 | 0.24 | 0.22 | 0.23 | 0.21 |
| | | | [3.10] | [2.85] | [2.67] | [2.43] | [2.54] | [2.34] |

**Table 3:** Conditional sort on size and SGPC

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

| | | SGPC Quintiles | | | | | SGPC Strategies | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (L) | (2) | (3) | (4) | (H) | $r^e$ | $\alpha_{CAPM}$ | $\alpha_{FF3}$ | $\alpha_{FF4}$ | $\alpha_{FF5}$ | $\alpha_{FF6}$ |
| Size quintiles | (1) | 0.40 | 0.75 | 0.99 | 0.93 | 0.89 | 0.49 | 0.48 | 0.52 | 0.41 | 0.59 | 0.50 |
| | | [1.63] | [2.89] | [3.86] | [3.67] | [3.30] | [4.73] | [4.67] | [5.28] | [4.21] | [6.15] | [5.32] |
| | (2) | 0.58 | 0.72 | 0.86 | 0.88 | 0.86 | 0.28 | 0.29 | 0.27 | 0.23 | 0.32 | 0.30 |
| | | [2.55] | [3.03] | [3.61] | [3.89] | [3.67] | [3.22] | [3.25] | [3.12] | [2.67] | [3.78] | [3.40] |
| | (3) | 0.57 | 0.67 | 0.79 | 0.83 | 0.89 | 0.32 | 0.32 | 0.32 | 0.30 | 0.33 | 0.31 |
| | | [2.76] | [2.98] | [3.46] | [3.93] | [4.30] | [4.39] | [4.30] | [4.30] | [3.99] | [4.29] | [4.08] |
| | (4) | 0.46 | 0.66 | 0.80 | 0.79 | 0.79 | 0.34 | 0.36 | 0.33 | 0.30 | 0.21 | 0.20 |
| | | [2.31] | [3.12] | [3.79] | [3.96] | [4.17] | [4.47] | [4.66] | [4.30] | [3.88] | [2.71] | [2.56] |
| | (5) | 0.43 | 0.48 | 0.54 | 0.53 | 0.71 | 0.28 | 0.26 | 0.26 | 0.23 | 0.25 | 0.23 |
| | | [2.53] | [2.52] | [2.95] | [3.05] | [4.21] | [2.99] | [2.84] | [2.77] | [2.40] | [2.69] | [2.44] |

Panel B: Portfolio average number of firms and market capitalization

| | | SGPC Quintiles<br>Average $n$ | | | | | SGPC Quintiles<br>Average market capitalization ($\$10^6$) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | (L) | (2) | (3) | (4) | (H) | (L) | (2) | (3) | (4) | (H) |
| Size quintiles | (1) | 389 | 390 | 388 | 387 | 387 | 32 | 34 | 39 | 28 | 30 |
| | (2) | 111 | 111 | 110 | 110 | 110 | 57 | 56 | 56 | 56 | 56 |
| | (3) | 81 | 81 | 80 | 80 | 81 | 98 | 96 | 98 | 100 | 99 |
| | (4) | 68 | 68 | 68 | 68 | 68 | 206 | 207 | 214 | 213 | 214 |
| | (5) | 62 | 62 | 62 | 62 | 62 | 1389 | 1419 | 1752 | 1626 | 1743 |

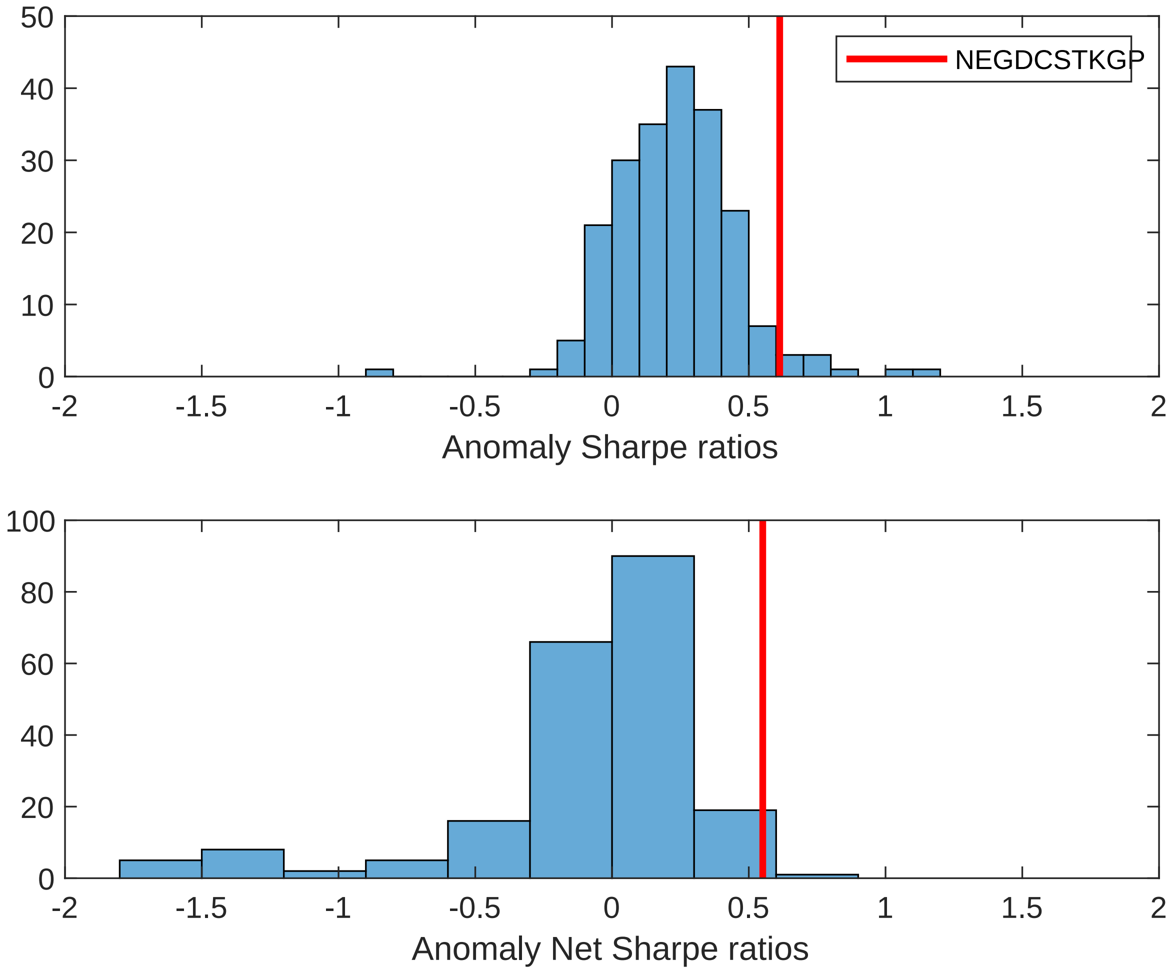


**Figure 2:** Distribution of Sharpe ratios.
This figure plots a histogram of Sharpe ratios for 212 anomalies, and compares the Sharpe ratio of the SGPC with them (red vertical line). Panel A plots results for gross Sharpe ratios. Panel B plots results for net Sharpe ratios.

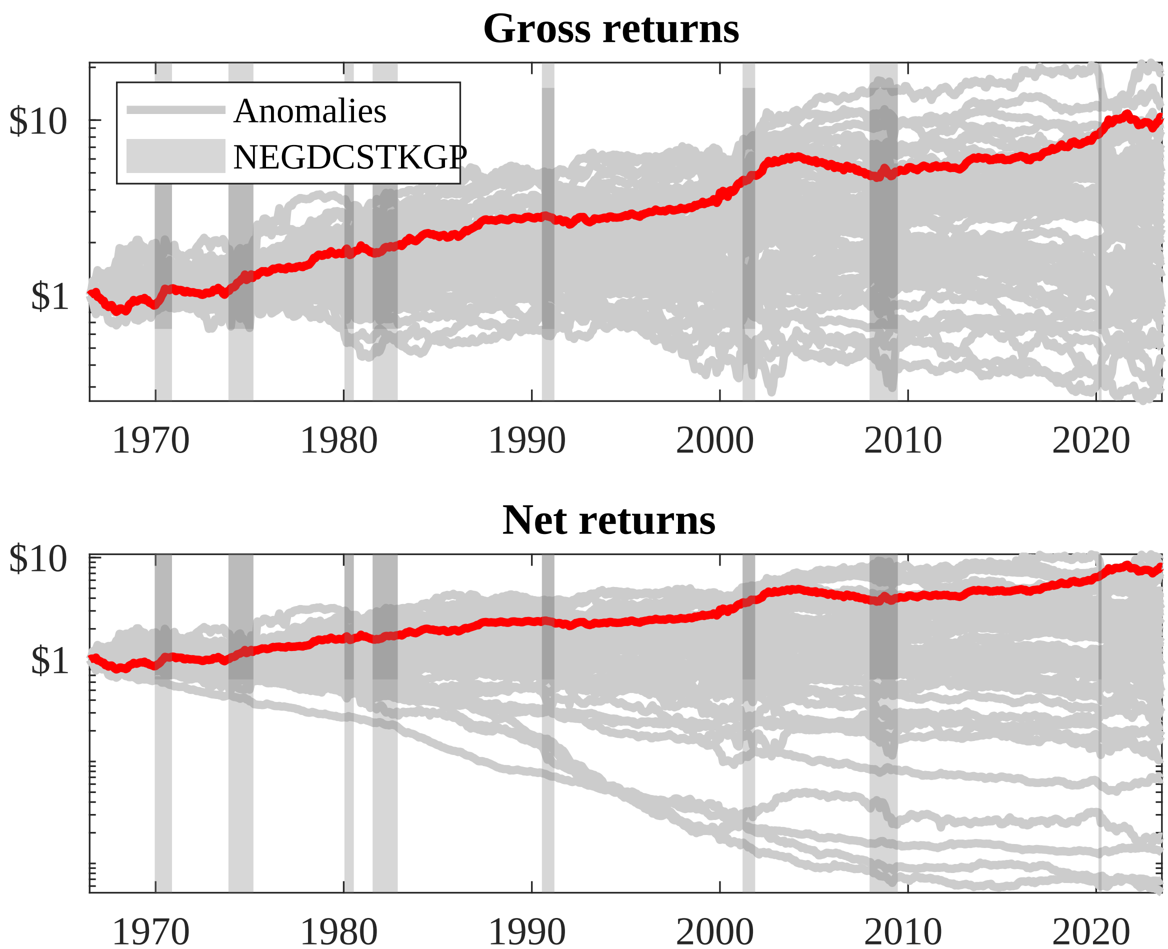


**Figure 3:** Dollar invested.
This figure plots the growth of a $1 invested in 212 anomaly trading strategies (gray lines), and compares those with the SGPC trading strategy (red line). The strategies are constructed using value-weighted quintile sorts using NYSE breakpoints. Panel A plots results for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

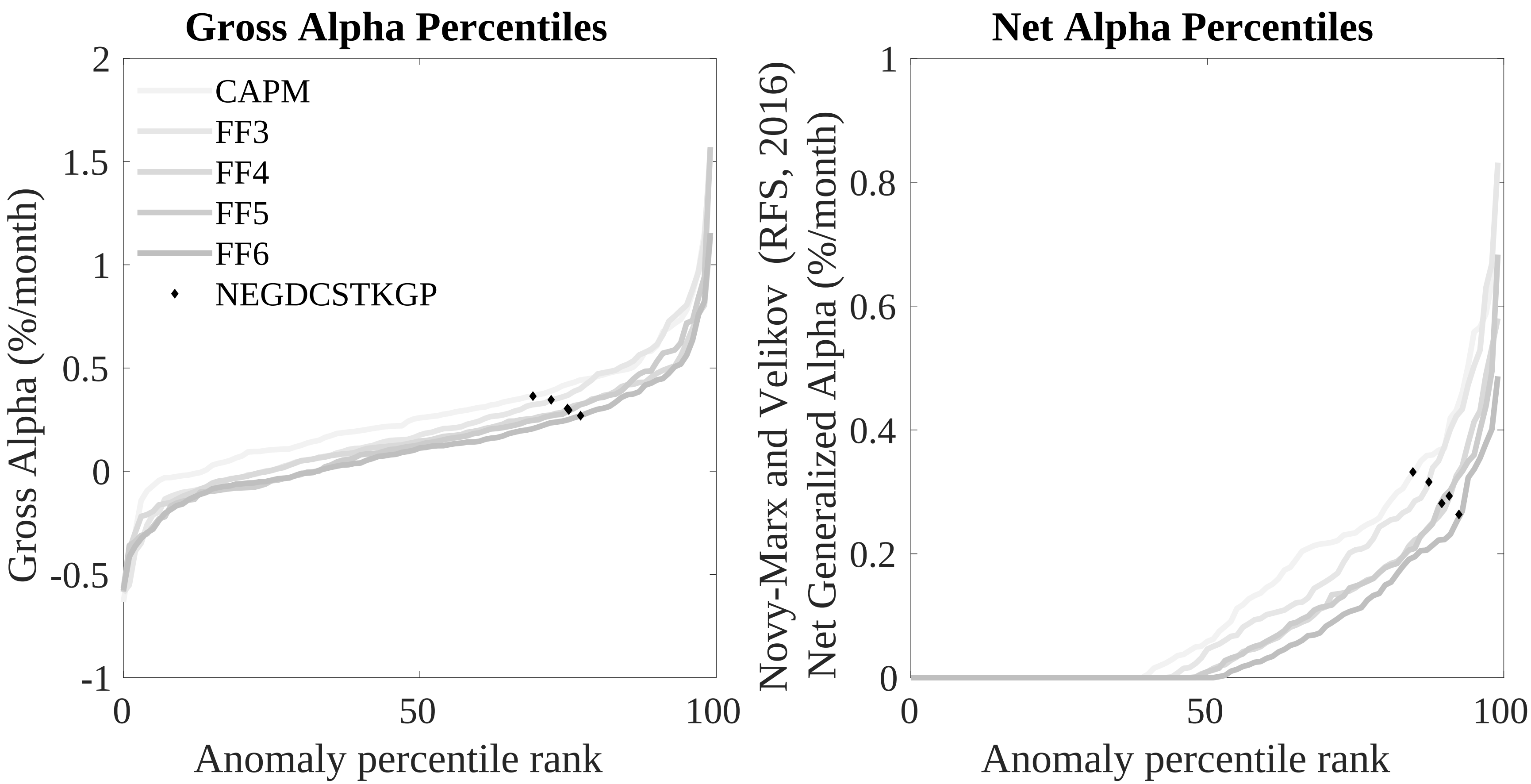


**Figure 4:** Gross and generalized net alpha percentiles of anomalies relative to factor models
This figure plots the percentile ranks for 212 anomaly trading strategies in terms of alphas (solid lines), and compares those with the SGPC trading strategy alphas (diamonds). The strategies are constructed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

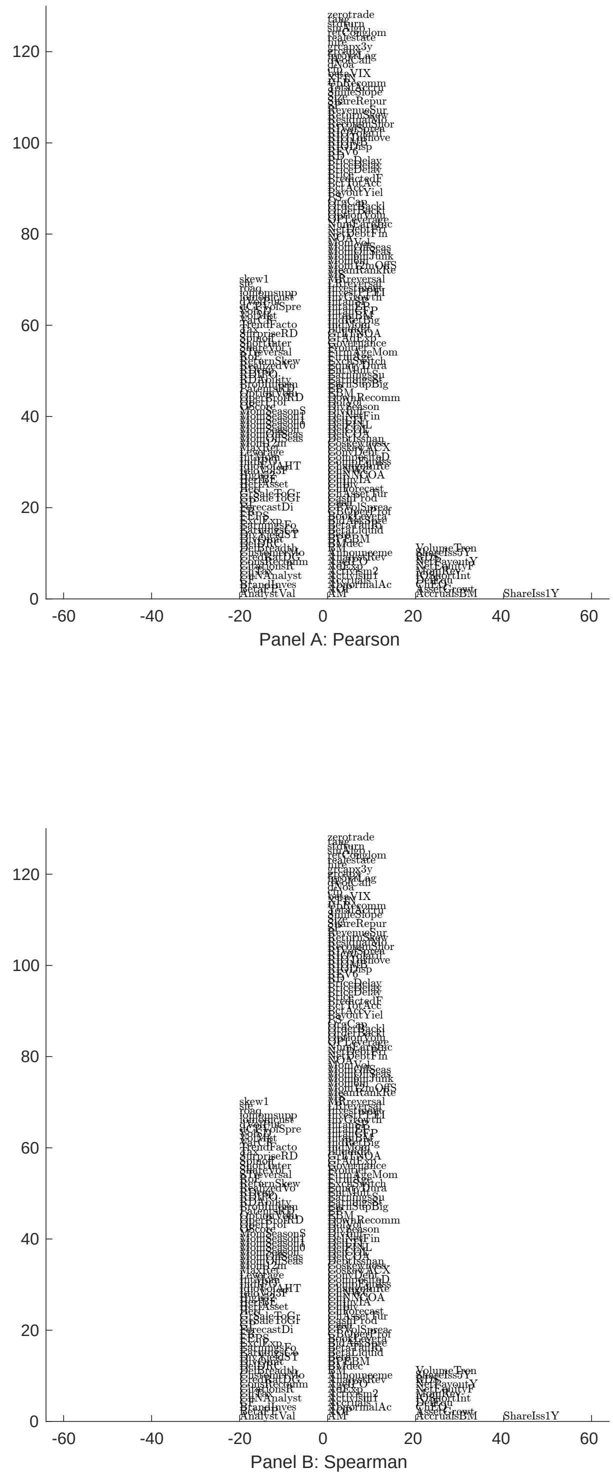


**Figure 5:** Distribution of correlations.
This figure plots a name histogram of correlations of 210 filtered anomaly signals with SGPC. The correlations are pooled. Panel A plots Pearson correlations, while Panel B plots Spearman rank correlations.

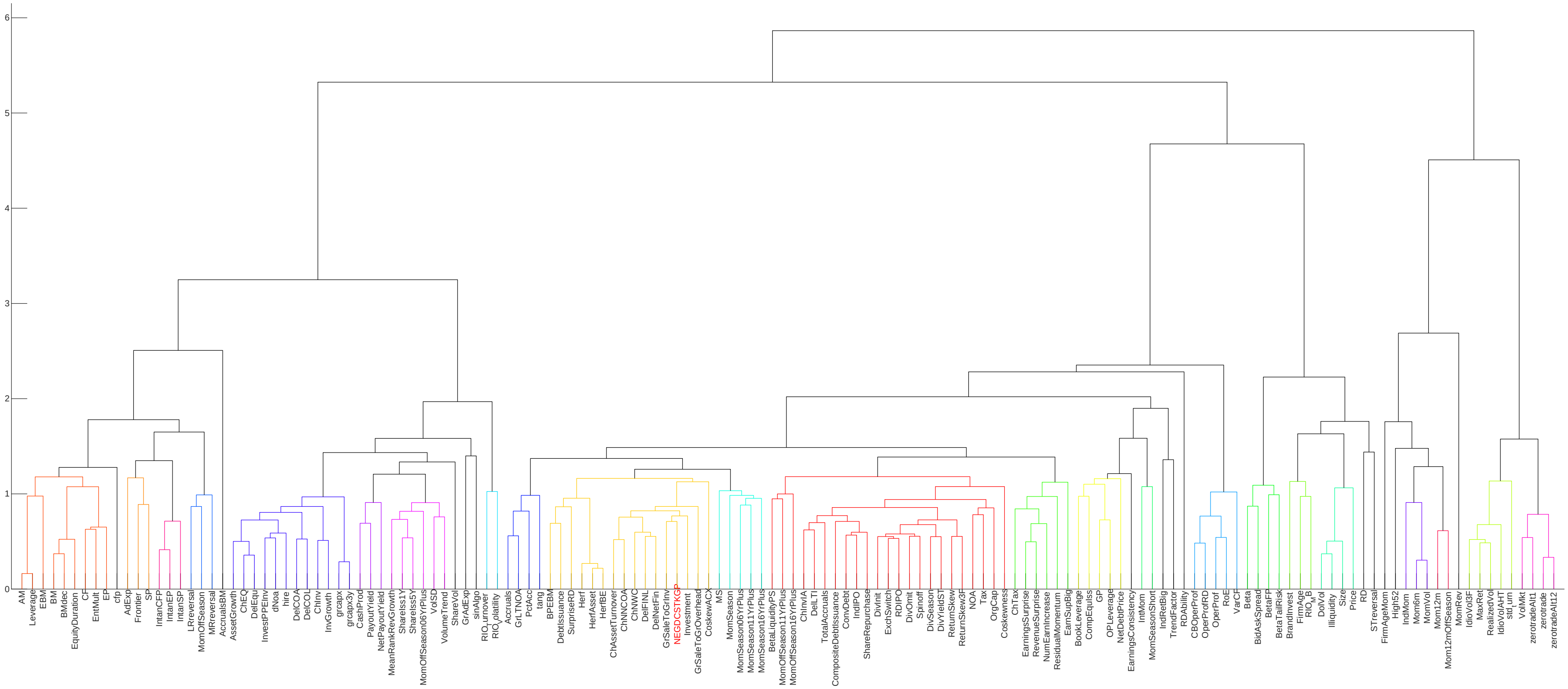

**Figure 6:** Agglomerative hierarchical cluster plot
This figure plots an agglomerative hierarchical cluster plot using Ward's minimum method and a maximum of 10 clusters.

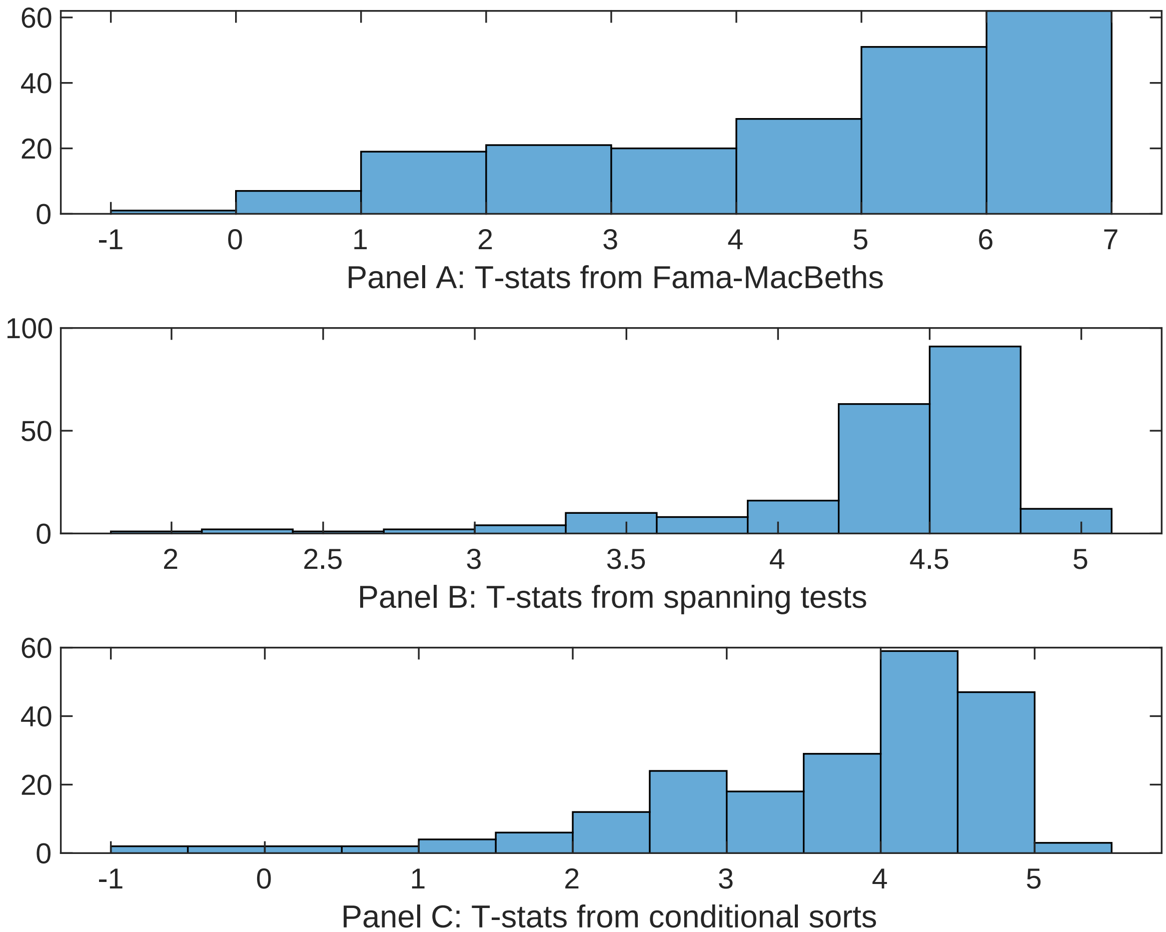


**Figure 7:** Distribution of t-stats on conditioning strategies
This figure plots histograms of t-statistics for predictability tests of SGPC conditioning on each of the 210 filtered anomaly signals one at a time. Panel A reports t-statistics on $\beta_{SGPC}$ from Fama-MacBeth regressions of the form $r_{i,t} = \alpha + \beta_{SGPC} SGPC_{i,t} + \beta_X X_{i,t} + \epsilon_{i,t}$, where $X$ stands for one of the 210 filtered anomaly signals at a time. Panel B plots t-statistics on $\alpha$ from spanning tests of the form: $r_{SGPC,t} = \alpha + \beta r_{X,t} + \epsilon_t$, 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 SGPC. Stocks are finally grouped into five SGPC portfolios by combining stocks within each anomaly sorting portfolio. The panel plots the t-statistics on the average returns of these conditional double-sorted SGPC 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 SGPC. and the six most closely related anomalies. The regressions take the following form: $r_{i,t} = \alpha + \beta_{SGPC} SGPC_{i,t} + \sum_{k=1}^{s} ix\beta_{X_k} X_{i,t}^k + \epsilon_{i,t}$. The six most closely related anomalies, $X$, are Growth in book equity, Share issuance (1 year), Change in equity to assets, Momentum and LT Reversal, Share issuance (5 year), Long-run reversal. 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 | 0.13 | 0.13 | 0.42 | 0.13 | 0.13 | 0.14 |
| | [7.40] | [5.70] | [5.64] | [1.23] | [6.06] | [5.79] | [3.24] |
| SGPC | 0.13 | 0.16 | 0.14 | 0.11 | 0.14 | 0.14 | 0.66 |
| | [4.83] | [5.76] | [5.46] | [0.97] | [5.02] | [5.61] | [0.43] |
| Anomaly 1 | 0.49 | | | | | | 0.67 |
| | [4.40] | | | | | | [2.63] |
| Anomaly 2 | | 0.25 | | | | | -0.14 |
| | | [5.53] | | | | | [-0.99] |
| Anomaly 3 | | | 0.15 | | | | -0.14 |
| | | | [4.21] | | | | [-1.74] |
| Anomaly 4 | | | | 0.11 | | | 0.87 |
| | | | | [4.27] | | | [2.67] |
| Anomaly 5 | | | | | 0.38 | | 0.44 |
| | | | | | [4.41] | | [1.45] |
| Anomaly 6 | | | | | | 0.27 | -0.28 |
| | | | | | | [2.88] | [-0.21] |
| # months | 684 | 679 | 684 | 636 | 679 | 679 | 606 |
| $\bar{R}^2(\%)$ | 0 | 0 | 0 | 2 | 0 | 1 | 0 |

**Table 5:** Spanning tests controlling for most closely related anomalies

This table presents spanning tests results of regressing returns to the SGPC trading strategy on trading strategies exploiting the six most closely related anomalies. The regressions take the following form: $r_t^{SGPC} = \alpha + \sum_{k=1}^{6} \beta_{X_k} r_t^{X_k} + \sum_{j=1}^{6} \beta_{f_j} r_t^{f_j} + \epsilon_t$, 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 Growth in book equity, Share issuance (1 year), Change in equity to assets, Momentum and LT Reversal, Share issuance (5 year), Long-run reversal. 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 | 0.27 | 0.25 | 0.29 | 0.24 | 0.24 | 0.24 | 0.23 |
| | [3.53] | [3.22] | [3.66] | [3.07] | [3.10] | [3.12] | [3.12] |
| Anomaly 1 | 34.76 | | | | | | 34.93 |
| | [8.28] | | | | | | [5.78] |
| Anomaly 2 | | 25.09 | | | | | 16.27 |
| | | [6.37] | | | | | [3.64] |
| Anomaly 3 | | | 18.76 | | | | -13.42 |
| | | | [4.54] | | | | [-2.39] |
| Anomaly 4 | | | | 3.71 | | | 3.37 |
| | | | | [3.72] | | | [3.32] |
| Anomaly 5 | | | | | 14.77 | | 4.55 |
| | | | | | [3.62] | | [1.07] |
| Anomaly 6 | | | | | | 6.73 | -0.88 |
| | | | | | | [2.92] | [-0.37] |
| mkt | 3.39 | 4.22 | 1.95 | 2.83 | 4.32 | 2.14 | 5.82 |
| | [1.90] | [2.33] | [1.06] | [1.54] | [2.28] | [1.16] | [3.18] |
| smb | 2.56 | 5.01 | 3.43 | 2.60 | 3.16 | 1.62 | 1.97 |
| | [0.99] | [1.92] | [1.29] | [0.96] | [1.18] | [0.57] | [0.72] |
| hml | -11.15 | -10.02 | -9.44 | -7.33 | -10.87 | -8.54 | -13.64 |
| | [-3.22] | [-2.83] | [-2.64] | [-2.05] | [-2.87] | [-2.33] | [-3.72] |
| rmw | -3.93 | -13.84 | -3.90 | -4.16 | -8.36 | -2.56 | -10.56 |
| | [-1.13] | [-3.69] | [-1.08] | [-1.16] | [-2.28] | [-0.69] | [-2.71] |
| cma | -6.20 | 16.50 | 8.77 | 25.98 | 24.22 | 24.89 | -4.02 |
| | [-0.95] | [2.97] | [1.30] | [4.86] | [4.42] | [4.50] | [-0.60] |
| umd | 3.93 | 4.09 | 4.86 | 0.86 | 4.51 | 5.06 | -0.06 |
| | [2.23] | [2.29] | [2.66] | [0.41] | [2.48] | [2.77] | [-0.03] |
| # months | 684 | 680 | 684 | 680 | 680 | 680 | 680 |
| $\bar{R}^2$(%) | 15 | 12 | 9 | 9 | 9 | 8 | 19 |

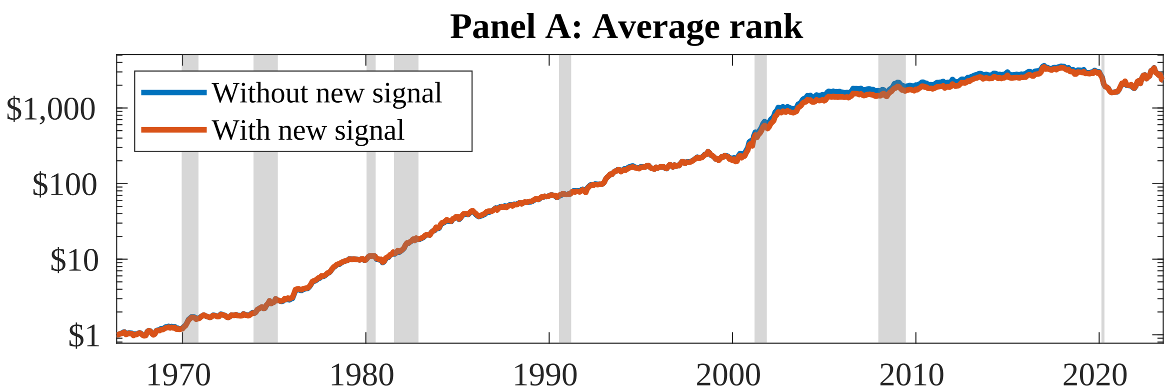


**Figure 8:** Combination strategy performance
This figure plots the growth of a $1 invested in trading strategies that combine multiple anomalies following Chen and Velikov (2022). In all panels, the blue solid lines indicate combination trading strategies that utilize 155 anomalies. The red solid lines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

# References


Ball, R., Brown, P., and Kruger, B. (2020). Predictable profitability and market underreaction. *Journal of Accounting Research*, 58(5):1021–1058.

Barberis, N. and Shleifer, A. (2003). Style investing. *Journal of Financial Economics*, 68(2):161–199.

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.

Cohen, L. and Frazzini, A. (2008). Economic links and predictable returns. *Journal of Finance*, 63(4):1977–2011.

Cooper, M. and Gulen, H. (2006). Is time-series-based predictability evident in real time? *Journal of Business*, 79(3):1263–1292.

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.

Hirshleifer, D., Li, J., and Yu, J. (2015). Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics*, 76:87–106.

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. (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*.