Directional Information in Equity Returns

Luca Del Viva [*], Carlo Sala and André B.M. Souza

Abstract

We document the existence of sign predictability in equity returns. An investment strategy that buys stocks deemed most likely to have positive returns and sells stocks with the lowest probability of positive returns generates about 1% monthly alpha and is not explained by established asset pricing models. The proposed strategy has higher Sharpe ratios and exhibits fewer crashes than the renowned momentum strategy. We show that profits from exploiting directional information are driven by shifts in retail investors' expectations after periods of excessive pessimism or optimism, rather than compensation for risk. We provide a simple model to motivate our findings.

JEL: G11, G12.

Keywords: Sign predictability, biased expectations, mispricing.

[*]Luca Del Viva **(corresponding author)**: ESADE Business School (Ramon Llull University), Barcelona, Spain. E-mail: luca.delviva@esade.edu. Carlo Sala: ESADE Business School (Ramon Llull University), Barcelona, Spain. E-mail: carlo.sala@esade.edu. André B.M. Souza: ESADE Business School (Ramon Llull University), Barcelona, Spain. E-mail: andre.souza@esade.edu. We thank Dante Amengual, Turan Bali, Irem Demirci, Pedro García Ares, Marco Grotteria, Giorgio Ottonello, Sheridan Titman, seminar participants at CUNEF, Cyprus University of Technology, NOVA Business School, XXIII Workshop on Quantitative Finance, FMARC 2022 and the 29^{th} Spanish Finance Forum. We gratefully acknowledge the financial support of Banco Sabadell, AGAUR-Generalitat de Catalunya-SGR 640 and MINECO project PID2021-126300NB-I00. **Declarations of interest: none.**

1 Introduction

Predicting future stock returns is an elusive goal (Malkiel and Fama (1970); Fama (1991)). Predicting the direction of change of returns, however, may be a feasible and profitable objective (Diebold and Christoffersen (2006); Diebold et al. (2007)). In this paper, we document the existence of directional predictability in equity returns and introduce the directional high-minus-low (D-HML) strategy: a trading strategy aimed at exploiting directional information in the time series and cross-section of stock returns. The proposed strategy sorts stocks according to forecasts of the conditional probability of future positive returns, buying (selling) stocks with the highest (lowest) predicted probability of positive returns. We carry out a comprehensive backtest of the newly proposed strategy and compare its performance with the well-known momentum strategy (Jegadeesh and Titman (1993)). The D-HML strategy generates an average raw return of 0.77% per month, exhibits positively skewed returns with a smaller probability of large losses when compared to the momentum strategy, and is not spanned by established asset pricing models that include the market, size, book-to-market, profitability, investment, momentum, short-term reversal, and long-term reversal factors. We provide a theoretical model that connects D-HML profits to revisions in investors' expectations following periods of excessive pessimism or optimism, and we document this connection empirically.

Earlier studies have documented the existence of directional predictability in returns. Breen et al. (1989) find that the one-month interest rate is useful in forecasting the signs of the excess market return. Leung et al. (2000) provide evidence that models aimed at predicting the directions of returns generate better directional forecasts than models that target the conditional mean of returns. Diebold and Christoffersen (2006) show that directional predictability is a consequence of persistent volatility dynamics, and hence to be expected even in the absence of conditional mean dependence. In a related strand of the literature, Maheu and McCurdy (2000) incorporate duration — the length, in months, of the current bull or bear market — into models for the conditional mean and variance of market returns, and find that the market returns in bull markets are a decreasing function of the duration of the current regime. Timmerman and Lunde (2004) find that the longer a bull market has lasted, the lower the probability that it terminates next period and, conversely, the longer a bear market has lasted, the higher the probability that it terminates next period. We build on both strands of the literature and construct probability forecasts based on lagged market returns, a measure of volatility, and the length of the current run of monthly, weekly, and daily returns of each stock.

Using a sample of U.S. equity returns spanning the period from August 1926 to December 2022, we document that our probability forecasts are informative about future return directions, thus presenting evidence of directional predictability in equity returns. We then study whether this directional predictability could have been exploited by market participants by inspecting the properties of the D-HML strategy. The D-HML strategy sorts the cross-section of probability forecasts into deciles and buys (sells) stocks in the top (bottom) decile. We find that stocks in the top decile of the predicted probabilities are about 7% more likely to have positive returns in the following period, have substantially higher average and median returns, and are more (less) likely to be future winners (losers) when compared to stocks in the bottom decile of the predicted probabilities.

To put our results into perspective, we compare the D-HML strategy to the widely used and notoriously profitable momentum strategy (WML). Despite being recognized to deliver high returns, momentum strategies are strongly affected by infrequent yet sizable crashes, a feature that renders such strategies unappealing for many risk-averse investors. As a consequence of these crashes, momentum returns are highly negatively skewed, and with a large excess kurtosis. In our sample, the momentum strategy generates an average raw return of 1.03% per month, with a volatility of 7.85, leading to an annualized Sharpe ratio of 0.34. In addition, WML returns have a skewness of -2.12 and a kurtosis of 21.24. In contrast, the D-HML strategy has an average raw return of 0.77% and a volatility of 3.24, leading to an annualized Sharpe ratio of 0.54. D-HML returns have a skewness of 1.02 and a kurtosis of 10.52. Moreover, the 5% quantile of WML returns corresponds, approximately, to the 0.1% quantile of D-HML returns, highlighting that large negative returns are substantially more likely for the WML portfolio than for the D-HML.

We carry out a series of tests to identify the source of the directional predictability and the D-HML return premium. We first examine whether profits from the D-HML strategy reflect compensation for higher levels of risk, as measured by a set of standard risk measures. The long leg of the D-HML portfolio is associated with lower risks when compared to the short leg, suggesting that profits from the D-HML strategy do not reflect compensation for bearing risk. We then evaluate whether the D-HML profits could be associated with conditional risk exposures, as measured by the Instrumented Principal Components (IPCA) approach of Kelly et al. (2021). We find that our probability forecasts remain informative in predicting returns after controlling for IPCA return predictions and that a D-HML portfolio formed on IPCA residuals generates a positive and significant alpha. Collectively, these findings imply that the profits from D-HML cannot be entirely attributed to compensation for risk.

We argue that the D-HML portfolio profits from revisions in investors' expectations following periods of excessive pessimism or optimism, in line with the biased expectations framework of Engelberg et al. (2018). In this framework, investors are excessively optimistic about some stocks and excessively pessimistic about others. Upon information arrival, investors' beliefs are updated and prices corrected. We build on this framework and introduce a theoretical model based on Barberis et al. (1998). In our model, firm earnings depend on the state of the market, which may be *good* or *bad*. Expected earnings are higher in the *good* state than in the *bad*. Investors receive random information shocks on the days prior to the earnings announcement and try to anticipate the state of the market. We model investors' beliefs as a Markov process in which a series of consecutive positive (negative) information shocks reinforces their beliefs about being in the *good* (*bad*) state. Our model generates both overpricing and underpricing. Overpricing attains when the investor observes a sequence of consecutive positive information shocks prior to earnings announcement, in which case they overestimate the probability of being in the *good* state. Conversely, underpricing attains when investors observe a sequence of consecutive negative information shocks prior to the earnings announcement, increasing their beliefs about being in the *bad* state.

Empirically, we document that stocks that experience a sequence of consecutive negative returns are both more likely to experience a positive return and tend to have a higher return on the month that follows when compared to their counterparts with sequences of consecutive positive returns. This effect is more pronounced for months in which firms release earnings announcements, suggesting that investors are overly pessimistic (or optimistic) in the build-up to earnings releases, in line with our theoretical model. We also document that the D-HML premium is significantly larger for stocks with lower institutional ownership, suggesting that retail investors are more likely to hold biased expectations.

Using either the Baker and Wurgler (2006), Huang et al. (2014), or the University of Michigan Consumer Sentiment indices as proxies for investor sentiment, we find that the short leg of the D-HML strategy is the main driver of profits during positive investor sentiment periods. In contrast, during negative investor sentiment periods, most of the profits stem from the long leg of the D-HML portfolio. This suggests that the D-HML strategy sells stocks for which investors are overly optimistic and buys stocks for which investors seem excessively pessimistic.

We further explore the role of the expectations revision in D-HML profits by using options implied skewness to proxy for investors' expectations about the distribution of future stock returns. We find that option investors significantly revise upwards (downwards) their expectations following the inclusion of a stock in the long (short) portfolios of the D-HML strategy, suggesting that the ex ante expectations were not aligned with the ex post realizations for the stocks selected in the D-HML portfolio.

We present a number of additional results. First, we document that the positive premium associated with the D-HML strategy cannot be explained by a set of established factors that are known to capture risk, mispricing, or behavioral biases. Second, we show that limits to arbitrage provide a partial explanation to the D-HML premium. Although D-HML profits are substantially larger for stocks with higher limits to arbitrage, the premium remains positive and significant for stocks with low limits to arbitrage. Finally, we document that ranks based on our probability forecasts are informative about future cross-sectional return ranks. Stocks in the top decile of the probability forecasts are ranked, on average, about 5.21 percentiles higher than those in the bottom decile, and this increase in ranking is associated with an average increase of 1.71% in one-month-ahead returns.

We carry out a number of robustness checks to assert the profitability of the proposed strategy. Our results are robust to the inclusion of transaction costs, the exclusion of penny stocks, different sample periods, different portfolio structures (equal- vs. value-weighted), and alternative directional forecasting models.

The remainder of this paper is organized as follows. Section 2 describes the sample and the construction of the probability forecasts. Section 3 examines the performance of the D-HML strategy. Section 4 tests risk-based explanations of the D-HML strategy. Section 5 explores the relation between investors' biased expectations and the D-HML premium. Section 6 presents a number of additional results and robustness checks. Section 7 concludes.

2 Data and Directional Predictability

2.1 Sample Description

Our data consist of monthly and daily stock prices for all NYSE/AMEX/NASDAQ ordinary equities (share codes 10 and 11) from the Center for Research in Security Prices (CRSP), covering the period from July 1926 to December 2022. We require each firm to have a minimum of one year of available stock prices, and we append delisting returns when available. Weekly returns are calculated by accumulating logarithmic daily returns over a week. Accounting data and quarterly earnings announcement dates are obtained from COMPUSTAT. Institutional ownership data are sourced from the Thomson–Reuters Institutional Holdings (13-F) database, with the sample period spanning from January 1980 to June 2022. Lastly, we complement our sample with stock-level daily option data from Ivy DB OptionMetrics, covering the period from January 1996 to December 2020.

2.2 Directional Predictability of Equity Returns

Denoting the return of stock i at time t as $r_{i,t}$, our objective is to forecast $\mathbb{P}_t(r_{i,t+1} > 0)$, where $\mathbb{P}_t(\cdot)$ represents the conditional probability measure based on the information available at time t. The dependent variable is constructed as $r^+_{i,t} = \mathbb{1}(r_{i,t} > 0)$, taking a value of one if the return of stock i in month t is strictly positive and zero otherwise. We consider several potential predictors identified in previous literature. We use the lengths of the current runs of daily, weekly, and monthly returns as measures of duration (Maheu and McCurdy (2000); Timmerman and Lunde (2004)). We argue that these duration measures capture excessive investors' optimism or pessimism on each stock. Volatility dynamics have also been shown (Diebold and Christoffersen (2006)) to predict return directions. Therefore, we include a measure of idiosyncratic variance in our model. Finally, we include lagged market returns to control for the current economic environment. All predictors considered are available at the time the forecast is made, and our estimates do not suffer from look-ahead bias. We obtain conditional probability forecasts from a linear probability model by estimating the following pooled regression:

$$
\begin{aligned}
r_{i,t}^{+} = \delta_0 + \delta_1 \, IV_{i,t-1} + \delta_2 \, r_{m,t-1} + \beta_{+}^{m} P_{i,t-1}^{m} + \beta_{+}^{w} P_{i,t-1}^{w} + \beta_{+}^{d} P_{i,t-1}^{d} + \\
\beta_{-}^{m} N_{i,t-1}^{m} + \beta_{-}^{w} N_{i,t-1}^{w} + \beta_{-}^{d} N_{i,t-1}^{d} + u_{i,t},
\end{aligned}
\tag{1}
$$

where $IV_{i,t-1}$ is the lagged idiosyncratic variance defined as the variance of the residuals after estimating the market model with daily returns over a month, $r_{m,t-1}$ is the lagged market return, and $\{P_{i,t-1}^{j}, N_{i,t-1}^{j}\}$ for $j \in \{m, w, d\}$ represent the duration of the current run of monthly, weekly, and daily positive and negative returns, respectively. We consider a maximum duration of 12 months, weeks or days.[1] We estimate Eq. (1) recursively with an expanding window, and out-of-sample forecasting starts in July 1932. We refer to the forecasts of $\mathbb{P}_t(r_{i,t+1} > 0)$ obtained from (1) as the probability score ($\mathrm{PS}_{i,t}$).[2]

Estimates of Eq. (1) based on the full sample indicate that longer runs are associated with the probability of future positive returns.[3] Runs of negative monthly, weekly, or daily returns are associated with an increase in the probability of a positive return in the following month. In contrast, runs of positive daily returns are associated with a decrease in the probability of a positive return, whereas runs of positive monthly and weekly returns are associated with a small increase in the probability of a positive return. Notably, the coefficients attached to daily runs are considerably larger than those for monthly and weekly runs. This suggests that a significant portion of the information contained in our probability forecasts is at-

[1] We stop at 12 months, weeks, or days as the frequency of observing a sequence of higher number periods with either consecutive positive or negative returns is close to zero.

[2] It is well-known that forecasts from linear probability models may be outside of the unit interval, and hence not valid probabilities (see, for instance, Greene (2017)). In our exercise, this happens for less than 0.01% of our forecasts.

[3] Table A.1 of the Online Appendix reports the coefficient estimates of Eq. (1) from the full sample, although our asset pricing tests are based on expanding window out-of-sample probability forecasts.

tributable to the cross-sectional variation in daily return sequences observed prior to portfolio formation.

To assess whether the predictive power remains in an out-of-sample setting, we evaluate the out-of-sample accuracy of our forecasts using standard tools from the market timing literature. Following Cumby and Modest (1997) and Breen et al. (1989), we test the significance of b in the following regressions:

$$r^{+}_{i,t+1} = a + b\mathbb{1}\{\text{PS}_{i,t} > 0.5\} + u_{i,t+1}, \quad \text{and} \tag{2}$$

$$r_{i,t+1} = a + b\mathbb{1}\{\text{PS}_{i,t} > 0.5\} + u_{i,t+1}. \tag{3}$$

Diebold and Christoffersen (2006) remark that such tests assess the quality of probability forecasts only through their implied directional forecast, treating forecasts of 0.4999 fundamentally different to those of 0.5001. We address this issue by testing for the significance of b in

$$r^{+}_{i,t+1} = a + b\text{PS}_{i,t} + u_{i,t+1}, \quad \text{and} \tag{4}$$

$$r_{i,t+1} = a + b\text{PS}_{i,t} + u_{i,t+1}. \tag{5}$$

A desirable feature of probability forecasts is that they help in predicting the cross-section of return directions. To assess this property, we also estimate Eqs. (2)–(5) using the Fama and MacBeth (1973) procedure.

[TABLE 1 ABOUT HERE]

Table 1 reports the results from this analysis. Models (1) and (2) in Table 1 show the estimated coefficients from Eqs. (2) and (4), whereas Models (3) and (4) present the results from Eqs. (3) and (5), respectively. Directional forecasts constructed from the probability score are significant in predicting both the direction and the conditional mean of future returns. OLS estimates contained in Models (1) and (3) of Table 1 indicate that a probability score greater than 0.5 is associated with a 4.51% increase in the probability of a positive monthly return and a 1.46% increase in average monthly returns. Moreover, the cross-sectional analysis indicates that stocks with PS greater than 0.5 are, on average, 5.23% more likely to have positive returns in a given month and exhibit, on average, 1.26% higher monthly returns than those with PS smaller than 0.5. Additionally, our probability forecasts, and not only their implied directions, are significant in predicting both the directions as well as the conditional means of returns. Models (2) and (4) in Table 1 show that a 1% increase in PS is associated with 0.86% increase in the probability of positive monthly returns, a 0.26% increase in average returns, and both coefficients are statistically significant. In the cross-section, a 1% increase in PS is associated with an approximately 1.02% increase in the probability of positive returns and a 0.23% increase in average returns.

It is well known that directional forecast accuracy is neither a necessary nor a sufficient condition for the existence of a profitable trading strategy that yields significant excess returns (Diebold and Lopez (1996); Satchell and Timmerman (1995); Cumby and Modest (1997)). In the next section, we construct a strategy based on the probability score and show that this strategy yields positive and significant excess and abnormal returns.

3 Directional Information and Equity Returns

3.1 Univariate Portfolio Sorting on Directional Forecasts

Building on the results from the previous section, for each month we sort stocks into deciles based on their probability score (PS). We then use the lagged market capitalization of each stock to create value-weighted portfolios for each month and decile. We label the strategy that takes a long position in stocks with high probability score (top decile) and a short position in stocks with low probability score (bottom decile) as the directional high-minus-low (D-HML) strategy.

[TABLE 2 ABOUT HERE]

Panel A of Table 2 presents the raw and abnormal returns of the D-HML portfolio. The abnormal returns (alphas) are estimated following the i) Sharpe (1964), Lintner (1965), and Mossin (1966) capital asset pricing model (CAPM); ii) Fama and French (1993) three-factor model (FF3); iii) Fama and French (2015) five-factor model (FF5); iv) FF5 augmented by the momentum; v) short- and long-term reversal factors (FF5+UMD+S/L-TR); vi) Hou et al. (2015, 2020b) and Hou et al. (2020a) five-factor model (Q5); vii) Q5 model augmented by the momentum and viii) the short- and long-term reversal factors (Q5+UMD+S/L-TR). The sample covers the period from July 1932 to December 2022 for the CAPM and the FF3 models, from July 1963 to December 2022 for the FF5 model, and from January 1967 to December 2022 for the Q5 model.

Panel A of Table 2 highlights that the monthly abnormal returns (alphas) of the D-HML portfolio range from 0.76% to 1.16%, depending on the sample period and asset pricing model considered. Over the sample period spanning from July 1932 to December 2022, the D-HML portfolio exhibited an annualized Sharpe ratio of 0.54 and a positive skewness of 1.02. Considering the more recent July 1963 (January 1967) to December 2022 periods, annualized Sharpe ratios are 0.65 (0.66), with skewness of 0.80 (0.77). The D-HML portfolio has positive and significant alpha spreads across all asset pricing models considered.

To put our results into perspective, we compare the portfolios built on our probability forecasts with the standard momentum winner-minus-loser (WML) strategy. We follow Jegadeesh and Titman (1993) and Fama and French (1996) and form portfolios by sorting stocks into deciles based on their momentum (MOM) variable, defined as the cumulative return from month $t-12$ to $t-2$ using NYSE breakpoints. For each month and decile, we build a value-weighted portfolio based on each stocks' lagged market capitalization, and we indicate as winner-minus-loser (WML) the strategy that buys past winners and sells past losers. Over the same sample period, the WML strategy generates an average monthly raw return of 1.03%, with monthly alphas ranging from 0.17% to 1.66%. Compared to the D-HML strategy, the WML strategy generates higher returns with a significantly higher volatility and lower skewness (-2.12 vs. 1.02), leading to an annualized Sharpe ratio that is about 60% of that observed for the D-HML strategy (0.34 vs. 0.54). Details and a discussion of the WML returns can be found in Section A and Table A.2 of the Online Appendix.

Panel B of Table 2 presents the abnormal return (alpha) for each decile portfolio constructed using the probability score over the period from July 1932 to December 2022, based on the three-factor model by Fama and French (1993). Abnormal returns increase monotonically from -0.43% to 0.39% per month as we move from the bottom to the top PS decile. Each decile portfolio has a market beta of approximately one and a near zero size beta. Returns from the low probability score portfolio have a positive and significant loading on the book-to-market factor, whereas returns from the high probability score portfolio have a negative loading on the same factor. The D-HML portfolio has insignificant market and size betas, and a significantly negative book-to-market beta. The maximum and minimum monthly returns of the D-HML portfolio are 28% and -17%, respectively. Panel C of Table 2 reports the abnormal return (alpha) for each decile portfolio built using the probability scored based on the sample periods and asset pricing models considered in Panel A. Abnormal returns from the low (high) probability score decile portfolios are negative (positive) across all asset pricing models and sample periods considered.

To verify whether sorts based on the probability score indeed capture stocks that are more likely to yield positive returns, we report, on Panel B of Table 2, the average percentage of positive returns in each portfolio over time for each decile of the probability score. On average, 46% of the stocks in the bottom PS decile have positive returns in the following month. In contrast, 53% of stocks in the top PS decile have, on average, positive returns. We also report the average percentage of stocks in each PS decile that are in the top (winners) and bottom (losers) deciles of one-month-ahead returns. This measure is informative about the probability of crashes in the overall portfolio, as crashes are typically associated with buying future losers and selling future winners. For the bottom PS decile, the average percentage of stocks that become winners is 10%, in contrast to 11% for the top PS decile. Moreover, the average number of future losers is 13% and 9% for those stocks in the bottom and top PS decile, respectively. We also report the average over time of the median market capitalization of the stocks in each portfolio. As a remark, we point out that the median market capitalization is evenly distributed across all portfolios constructed.

3.2 Crashes and Long-Term Performance

It is well-documented that, while generating high returns, the momentum strategy has heavy left tails, producing the so-called momentum crashes (Daniel and Moskowitz (2016)). To assess whether the proposed strategy is susceptible to similar crashes, we analyze the tails of the distribution of returns for the D-HML strategy and compare it with those of the WML strategy. Defining a crash as an episode in which the three-month cumulative excess returns of a strategy lies below momentum's 5% quantile (-16.20%), the WML portfolio experienced 55 crashes over our sample period, in stark contrast to the 4 crashes experienced by the D-HML strategy.

[FIGURE 1 ABOUT HERE]

Figure 1 displays a scatter plot of the cumulative three-month excess returns of the D-HML and WML strategies conditional on a crash as previously defined.[4] Points located above the 45-degree line represent a greater severity of crashes in the WML strategy when compared to the D-HML strategy. Remarkably, the WML strategy incurred larger losses than the D-HML strategy in all but one crash episode, in which both strategies experienced losses of similar magnitudes. The D-HML strategy experienced positive cumulative returns in about 69% of the momentum crashes. To gauge the impact of crashes on the long-term performance of both strategies, Figure 2 illustrates the cumulative returns of the WML and D-HML portfolios for the entire sample period and across three roughly 30-year sub-periods.

[FIGURE 2 ABOUT HERE]

As shown in Panel A of Figure 2, the WML portfolio is more volatile and experiences more crashes than the D-HML portfolio. In contrast, the D-HML strategy provides stable, albeit smaller, gains. Over the full sample, the consistent gains and the absence of large crashes leads to better long-term performance of the D-HML strategy when compared to the WML strategy. Panels B to D of Figure 2 plot the cumulative return over three roughly 30-year sub-periods. The WML strategy performed better than the D-HML only in the sub-sample from January 1961 to December 1990. In the periods from July 1932 to December 1960 (Panel B) and from January 1991 to December 2022 (Panel D), the D-HML strategy outperformed the WML. We also remark that the WML returns become highly negative for the period starting in July 1932, whereas the D-HML returns remain positive throughout the sample.

Finally, we investigate the performance of the D-HML and WML strategies across different market states. As in Daniel and Moskowitz (2016), we find that the

[4] Figure A.1 in the Online Appendix presents results from conditioning on a D-HML, instead of a WML, crash.

WML strategy delivers a positive and significant alpha spread during bull markets of 1.41% per month. During bear markets, however, the WML strategy delivers a negative alpha spread of -0.33% per month. The D-HML strategy, on the other hand, delivers positive and significant abnormal returns regardless of the market state, with an alpha spread of 0.66% per month during bull markets. During bear markets, the D-HML has a 0.69% per month premium, delivering an alpha spread of 1.35% per month. Additionally, the WML strategy performs poorly during rebounds following bear markets, whereas the D-HML strategy does not seem to be affected by market rebounds. Section B and Tables A.3 and A.4 in the Online Appendix contain a detailed analysis of both the WML and the D-HML strategies across different market states.

Overall, the D-HML strategy exhibits consistent positive returns across different market states, fewer crashes, and superior long-term performance when compared to the WML strategy. The D-HML strategy's outperformance is particularly pronounced over the more recent period spanning from January 1991 to December 2022.

4 Risk and the D-HML Premium

4.1 Firm Characteristics

We investigate whether the positive D-HML premium arises as compensation for risk. We start by examining whether sorts based on the probability score (PS) are associated with a set of standard risk measures that are based on stock characteristics. Using two months of daily data, we construct, for each stock, the market beta (Market Beta), market beta with respect to negative and positive market returns (Neg. Beta and Pos. Beta), total volatility of returns (Total Vol.), downside volatility (Neg. Vol.), upside volatility (Pos. Vol.), volatility of the CAPM residual or idiosyncratic returns (Idio Vol.), downside volatility of idiosyncratic returns (Neg.

Idio. Vol.), and upside volatility of idiosyncratic returns (Pos. Idio. Vol.).

[TABLE 3 ABOUT HERE]

Table 3 reports the average of each risk measure computed in the pre-formation, holding and post-formation periods. The risk measures are computed over two months of daily data. Portfolios are formed at month t and kept until month $t+1$. The pre-formation period corresponds to the two-month period from $t-1$ to t. The holding period covers from month t to month $t+1$, and the post-formation period runs from month $t+1$ to $t+2$. Stocks in the high PS portfolio are associated with lower market beta relative to their low PS counterparts in the period leading to portfolio formation and during the holding period, with a similar market beta in the post-formation period. This suggests that the exposure to market risk is not the source of the D-HML premium. Moreover, stocks in the high PS portfolio have lower total and idiosyncratic volatility relative to their low PS counterparts.

In addition to the risk measures reported in Table 3, we also compare stocks in the high and low PS portfolios according to a number of firm characteristics. We consider the characteristics employed in Freyberger et al. (2020). Details on the construction of the characteristics can be found in Section C of the Online Appendix, and the results are reported in Table A.5 of the Online Appendix. Overall, firms with low and high probability scores show small differences across several characteristics. The most notable differences are that firms in the top decile of PS have higher earnings-to-price ratio (E2P), higher return on assets (ROA), higher long-term reversal (LTR), lower previous-month returns (STR), lower capital intensity (D2A), and lower sales-to-market capitalization ratio (S2P). Overall, these differences seem insufficient to justify the higher premium associated with high PS stocks. Taken together, our results provide evidence against the hypothesis that the positive D-HML premium reflects compensation for bearing risk.

4.2 Conditional Risk Exposures and the D-HML Premium

Kelly et al. (2021) highlight the limitations of traditional risk measures and employ the Instrumented Principal Components Analysis (IPCA) of Kelly et al. (2019) constructed with 36 firm-level observable characteristics to estimate a model that captures conditional risk exposures. Kelly et al. (2021) show that, once such estimates of the conditional expected returns are accounted for, the premium associated with the momentum and residual momentum (Grundy and Martin (2015)) strategies becomes insignificant, suggesting that profits from these strategies reflect compensation for conditional risk exposures.

We perform a similar analysis and test whether the positive premium associated with the D-HML strategy can be explained by IPCA-based estimates of conditional expected returns. If the return premium associated with the probability score is indeed a consequence of higher risk exposures, controlling for the IPCA return predictions should render this premium insignificant. On the other hand, should the D-HML premium remain significantly positive once the IPCA-based conditional expectations are accounted for, this would present evidence against the hypothesis that the D-HML premium reflects conditional risk exposures as captured by the IPCA model. As in Kelly et al. (2021), we estimate the IPCA using the 36 firm-level characteristics defined in Freyberger et al. (2020).[5]

We follow Kelly et al. (2021) and run the following multivariate panel regression:

$$r_{i,t+1} = c_0 + c_1 \mathrm{MOM}_{i,t} + c_2 \mathrm{IPCA}_{i,t} + c_3 \mathrm{PS}_{i,t} + u_{i,t+1}, \tag{6}$$

where $\mathrm{IPCA}_{i,t}$ are IPCA based expected returns. This regression allows us to assess whether the information content of a given variable is nested by that of other variables. Following Kelly et al. (2021), we estimate joint predictive regressions

[5] Codes for the IPCA estimation have been downloaded from Seth Pruitt's website (https://sethpruitt.net/research/downloads/).

with the raw signals (MOM, IPCA and PS), as well as cross-sectional rankings generated from these variables.

[TABLE 4 ABOUT HERE]

The results from Model (1) in Panel A of Table 4 are aligned with those in Kelly et al. (2021) and highlight that the information content in raw momentum is nested by that of the IPCA. In addition, the results from Model (2) show that our probability forecasts also nest the information content of raw momentum. In the left panel, the results from Model (3) provide evidence that the information content of PS is complementary to that of the IPCA. This finding remains unchanged when raw momentum is included in the regression, as can be seen from Model (4)'s results. The results for signal ranks (Models (5) to (8)) are similar, with the exception that the ranked momentum has a positive and significant coefficient after controlling for the ranked PS.

Finally, to assess how much of the D-HML premium is explained by the IPCA, we carry out a univariate portfolio analysis using residual returns. In particular, for each firm i and month t, we calculate the residual return as the difference between actual stock returns and the conditional expected returns obtained from the IPCA model, that is, $\epsilon_{i,t} = r_{i,t} - \text{IPCA}_{i,t-1}$. We then build portfolios based on the probability score by value weighting the residual returns $\epsilon_{i,t}$. Panel B in Table 4 contains a similar analysis as the one contained in Panel A of Table 2 using the residual returns rather than actual returns. The results in Panel B in Table 4 indicate that the D-HML strategy built on IPCA residuals still generates significantly positive abnormal returns, providing further evidence that the D-HML premium cannot be fully explained by the IPCA-based conditional risk exposures.

Taken together, the results from the risk measures presented in Section 4.1 and the IPCA-based conditional risk measures discussed in this section suggest that the

D-HML profits cannot be attributed exclusively to compensation for risk.

5 Mispricing and the D-HML Premium

5.1 Investors' Biased Expectations

Following Engelberg et al. (2018), we investigate whether the D-HML premium can be attributed to mispricing arising from investors' biased expectations. Under this framework, investors hold biased expectations about the future cash flows of companies. These expectations are adjusted following information arrivals, which in turn leads to price corrections. We introduce a theoretical model based on Barberis et al. (1998) in which sequences of positive or negative information shocks generate mispricing through investors' biased expectations.

In our model, the distribution of firm earnings depends on the state of the market, a latent random variable. We assume there are two possible states, *good* and *bad*. Expected earnings are higher in the *good* state than in the *bad* state. Investors know that there are two possible states, but lack knowledge about the probabilities associated with each state. To correctly price the asset, investors must use a sequence of (random) observable information shocks to estimate the probability of being in each state. Shocks are observed sequentially in each period before earnings are paid out and the state revealed. The information shocks are random but may be related to the state of the world, that is the probability of a positive shock may depend on the probability of being in the *good* state. We model investors' expectations as a belief-reinforcing Markov process, in which a series of consecutive positive news reinforces investors' beliefs about being in the *good* state, and a series of consecutive bad news reinforces their belief of being in the *bad* state. In our model, overpricing (underpricing) attains when investors observe a sequence of consecutive positive (negative) information shocks prior to earnings announcement, in which case they overestimate (underestimate) the probability of being in the *good* state. Upon information arrival, the state is revealed and prices revert to fundamentals, correcting the mispricing arising from biased expectations. Profits from trading against biased expectations are therefore proportional to the price corrections following information arrivals, which, in turn, depend on the runs of positive or negative information shocks. Section A in the Appendix provides details on the model.

If the probability score as estimated in Eq. (1) correlates with investors' biased expectations, then the D-HML strategy profits from correctly identifying stocks for which investors hold overly pessimistic or optimistic expectations. One condition for the D-HML to profit from biased expectations is that it must be a contrarian strategy, that is, it must trade against investor expectations. We establish that this is indeed the case. The median stock in the long leg of the D-HML has had negative returns over the last two months, the last week, and the last two days prior to portfolio formation. In contrast, the median stock in the short leg of the D-HML has had positive returns on the last month, the last week, and the last three days prior to portfolio formation. Because the D-HML strategy commands a positive premium, investors held misaligned expectations for some stocks in the D-HML portfolio. In fact, we find that the average excess return for stocks in the long leg of the D-HML portfolio is 1.94% per month, whereas the average excess return for stocks in the short leg of the D-HML is -0.18% per month. In the month following the portfolio holding period, stocks in the long and short legs of the D-HML have similar returns (0.79% and 0.99% per month, respectively). This suggests that the D-HML trades on short-lived market inefficiencies that are quickly corrected upon information arrival, in line with the biased expectations hypothesis.[6]

If, as we argue, information arrivals and their associated price corrections are indeed related to directional return predictability and the D-HML premium, we

[6]Figure A.2 in the Online Appendix shows that the D-HML premium weakens at $t+2$ and remains statistically insignificant from month $t+3$ up to 24 months after the portfolio formation.

should also observe higher (lower) returns for stocks in the top (bottom) of our probability forecasts with earnings announcements in the portfolio holding period.

[FIGURE 3 ABOUT HERE]

Panel A of Figure 3 shows the average return for stocks in each decile of our probability forecasts, categorized into those with earnings announcements during the portfolio holding period and those without. Stocks in the top decile of our probability forecasts earn, on average, 2.19% in earning months, 18% higher than the 1.85% average returns in non-earning months. In contrast, stocks in the bottom decile of our probability forecasts earn, on average, -0.19% in earning months, 5% lower than the -0.18% average returns in non-earning months.[7] In addition, we find a significant increase in the length of runs of consecutive daily returns prior to earnings announcement days, followed by a significant decrease in the average run length after earnings announcements.[8] This provides further evidence that runs of consecutive returns are related to the build up and correction of investors' biased expectations around earnings announcements, in line with the cyclical biased expectations hypothesis (Linnainmaa and Zhang (2022)). It is worth noting that stocks in the top PS decile earn higher returns than those in the bottom PS decile both in months with earnings announcements and in those without. This could be attributed to the fact that earnings announcements are but one facet of information arrival. Numerous other types of information exist that can trigger price corrections in non-earnings announcement months, as emphasized by Engelberg et al. (2018).

To formally gauge the return premium in earnings announcement months, we regress excess daily returns on a dummy variable equal to 1 if a stock had earnings

[7]These results are based on the subset of stocks with earnings announcement data and, therefore, may differ from the full sample analysis.

[8]Figure A.3 in the Online Appendix contains the average length (runs) of consecutive days of negative and positive returns and the average length of the unsigned runs on a 30-day window around earnings announcements.

announcements in a given month (Earnings Month), the probability score (PS), and the interaction between the probability score and earnings announcements (PS $\times$ Earnings Month). In addition, we also include the *Net* measure of Engelberg et al. (2018) and its interaction with earnings announcement months. The *Net* measure is the difference between the number of anomalies for which a stock is classified in the long leg, minus the number of anomalies for which it is classified in the short leg.[9]

[TABLE 5 ABOUT HERE]

Table 5 reports the corresponding regression results. All regressions include month fixed effects. We include three lags of volatility and excess returns as control variables. Models (1) and (4) of Table 5 show that both our probability forecasts and an earnings-announcement dummy are associated with higher average excess returns. Moreover, the interaction between the probability forecasts and the earnings-announcement dummy is positive and significant. A one standard deviation increase in PS is associated with about a 0.23% additional excess returns in earnings-announcement months. Models (2) and (5) show that the *Net* measure is also associated with higher average excess returns and exhibits a similar pattern to the probability score. Models (3) and (6) report results from a regression that includes both the PS and the *Net* measure. All coefficients remain significant (at the 5% significance level) and with similar magnitudes, suggesting that the *Net* and the PS capture complementary aspects of mispricing due to investors' biased expectations.

If the D-HML portfolio indeed profits from mispricing, we would expect stocks with a high PS to show higher returns when there is more evidence of mispricing.

[9] We use a set of 171 known market anomalies from Chen and Zimmermann (2022). The full list of anomalies can be found in Table A.6 of the Online Appendix. All anomalies are rotated in order to deliver a positive return premium. Therefore, the top (bottom) decile of each anomaly represents underpriced (overpriced) stocks.

To verify this hypothesis, we construct tercile portfolios based on the *Net* and PS individually and compare these portfolios with those based on the intersection of stocks that are present in the top (bottom) 33% of both measures.

[TABLE 6 ABOUT HERE]

Table 6 reports the abnormal returns for the bottom and top 33% portfolios built on *Net*, PS, and the intersection of the two (PS $\cap$ *Net*). High minus low portfolios based exclusively on the *Net* generate monthly abnormal returns ranging from 0.26% (t-stat = 2.55) to 0.71% (t-stat = 8.90), depending on the asset pricing model considered. Analogously, the high minus low portfolios based exclusively on PS generate significant monthly abnormal returns ranging from 0.44% (t-stat = 5.74) to 0.60% (t-stat = 5.28). In line with the hypothesis that profits are higher for stocks that exhibit more evidence of mispricing, the abnormal returns from portfolios constructed based on the subset of stocks for which both measures agree on is substantially higher than those obtaining by individual sorts. Indeed, the abnormal returns range between 0.76% (t-stat=4.78) to 1.21% (t-stat=11.02), suggesting that PS profits are higher for stocks that exhibit more evidence of mispricing as measured by the *Net*.

Clearly, the *Net* is just one of the many measures of mispricing. To verify that the PS is associated with a general notion of mispricing, beyond what is measured by the *Net*, we value whether the ex ante PS is associated with an ex post mispricing measure. In particular, we compute $u_{i\,t} = \tilde{r}_{i\,t} - IPCA_{i\,t|t-1}$, where $\tilde{r}_{i\,t}$ denotes returns in excess of the risk-free rate and $IPCA_{i\,t|t-1}$ is the IPCA prediction for $\tilde{r}_{i\,t}$, obtained at time $t-1$. In this setting, $u_{i\,t}$ represents mispricing, as it captures the difference between the expected return for asset i based on conditional risk measures and its realized returns. If realized returns are higher than the predictions, the stock was undervalued. In contrast, if realized returns are lower than predicted, the stock was overvalued. As a remark, note that this is an ex post measure: it relies on realized returns that were not available at time $t-1$.

[TABLE 7 ABOUT HERE]

Panel A of Table 7 reports the average IPCA residual across probability score deciles. Each month, stocks are sorted into 10 deciles based on the probability score, and we compute the average IPCA residuals in each decile. Panel A of Table 7 shows that the average IPCA residuals increase monotonically from -0.69 to 0.62 as we move from the first to the last decile of the probability score. As negative (positive) IPCA residuals indicate the presence of overvaluation (undervaluation), the results contained in Panel A of Table 7 indicate that stocks in the bottom decile of PS were overvalued, whereas those on the top decile of PS were undervalued.

To further corroborate our results, we investigate the performance of PS returns across quintiles of lagged IPCA residuals. Despite being weaker instruments for mispricing, lagged IPCA residuals can be constructed ex ante. If the PS is associated with mispricing, the D-HML premium should be higher for stocks that are ex ante on the bottom and top quintiles of mispricing. The last column of Panel B of Table 7 shows that the D-HML premium is larger for stocks that feature in the top and bottom deciles of the mispricing proxy. Moreover, the D-HML premium is smaller and only marginally significant for stocks that are more accurately priced.

5.2 Biased Expectations and Institutional Ownership

We investigate whether the D-HML premium is larger for stocks predominantly owned by retail investors, as would be the case if retail investors were more susceptible to holding biased expectations than institutional investors. We use the percentage of institutional ownership from Thomson–Reuters Institutional Holdings Database (13F) in order to measure the degree of institutional ownership.[10] For this analysis, we focus exclusively on the subset of firms that are present both in the 13-F dataset and in our sample. To assess whether returns are higher in stocks with high retail participation, for every month we independently double sort stocks into five equally spaced quintiles based on the percentage of institutional ownership observed at the end of the previous quarter and on our probability score PS. We then form 25 value-weighted portfolios from the intersection of the two measures. Portfolios are rebalanced monthly. Panel A of Table 8 contains the abnormal returns (alphas) from the Fama and French (2015) five-factor model augmented with momentum (UMD), short-term reversal (STR), and long-term reversal (LTR).

[TABLE 8 ABOUT HERE]

Panel A of Table 8 shows that the D-HML premium decreases as we move from stocks with low institutional ownership (mostly held by retail investors) to stocks mostly held by institutional investors. Overall, our results indicate that the return premium associated with the probability score is higher for stocks with a larger presence of retail investors. This suggests that retail investors are more likely to hold biased expectations than institutional investors.

To complement our analysis, we also consider abnormal returns (alphas) from the low PS, high PS, and the D-HML portfolios obtained as the difference between low (bottom 33%) and high (top 33%) institutional ownership. We consider three measures of institutional ownership: (i) the total reported institutional ownership over the last quarter, (ii) the percentage of the institutional ownership, and (iii) the residual institutional ownership estimated as in Nagel (2005). Panel B of Table 8 reports our results. The D-HML portfolio profits obtained on stocks largely held by retail investors are more than double those obtained on stocks largely held by

[10]Where applicable, we complement this data with the shares outstanding obtained from CRSP.

institutional investors, when institutional investor ownership is measured as a percentage of overall shares outstanding, and this difference is statistically significant at the 1% significance level. Although varying in magnitude, the outcomes from the remaining two measures of institutional ownership corroborate these findings.

Collectively, our results indicate that the D-HML strategy's profits are more pronounced for stocks predominantly held by retail investors and that such investors are more susceptible to hold biased expectations.

5.3 Investor Sentiment and the D-HML Premium

Stambaugh et al. (2012) and Antoniou et al. (2016) document that overpricing is most prevalent during high-sentiment periods, when unsophisticated investors tend to be overly optimistic and more likely to participate in the market. We document that during low-sentiment periods, most of the D-HML profit stems from the long portfolio. In contrast, in high-sentiment periods, the short leg is responsible for the majority of profits. We measure investor sentiment using one of following three proxies: i) the Baker and Wurgler (2006) sentiment index, ii) the University of Michigan Consumer Sentiment Index, iii) the Huang et al. (2014) news-based sentiment index. For each of the sentiment indices considered, we build two binary variables taking the value of one if the previous month's sentiment index is above (below) its median, computed over the whole sample, and zero otherwise. We then run a regression of excess returns on the low and high PS portfolios against a set of risk factors and the sentiment dummies.

[TABLE 9 ABOUT HERE]

Table 9 reports the coefficients estimated with the Fama and French (2015) five-factor model augmented with the momentum (UMD), the short-term reversal (STR), and long-term reversal (LTR) factors. In line with Stambaugh et al. (2012)

and Antoniou et al. (2016), we find that the D-HML premium is higher in high-sentiment periods, in which investors tend to be excessively optimistic. This high sentiment premium is in large part driven by a substantial increase in the performance of the short leg of the D-HML portfolio. In addition, Panel B of Figure 3 (left plot) shows that the average return for stocks in the short leg of the D-HML is lower in earning announcement months but lower still in high-sentiment periods. In contrast, the right plot of Panel B of Figure 3 indicates that the average returns for stocks in the long leg of the D-HML is higher in earnings announcement months and higher still in low-sentiment periods. Our findings suggest that investors are overly optimistic in high-sentiment periods, and the D-HML strategy banks on this excessive optimism by short selling stocks for which investors seem excessively optimistic. On the other hand, in low-sentiment periods, investors tend to be overly pessimistic, and the D-HML strategy profits by buying stocks for which investors seem excessively pessimistic.

5.4 Probability Score and Options-Implied Skewness

As a final exercise, we use option-implied skewness as a forward-looking measure of expectations about jumps in the underlying stock price. Following Bali et al. (2019), we construct a non-parametric option-implied indicator based on the difference between monthly out-of-the-money (OTM) call and put options observed on the last day of the previous month.[11] See Section D of the Online Appendix for details on the construction of the implied skewness measures.

[FIGURE 4 ABOUT HERE]

Panel A of Figure 4 reports the average implied skewness of the top and bottom decile portfolios based on the probability score in the months around portfolio

[11]As a robustness check, we repeat the same analysis using the non-parametric approach proposed by Xing et al. (2010). Unreported results confirm the same findings.

formation, along with 95% (pointwise) confidence bands. Option investors have similar expectations regarding stocks in the top and bottom deciles of our probability score in the months leading to portfolio formation. During the holding period, stocks in the top decile of the probability score show a sizable increase in implied skewness, indicating that investors became more optimistic toward these stocks after they were included in the D-HML portfolio. This suggests that these are stocks for which investors may have received unexpected positive news. In contrast, stocks in the bottom decile of the probability score show a sizable decrease in implied skewness, which indicates that investors turned pessimistic toward these stocks post portfolio formation, suggesting that these stocks may have received unexpected negative news. Panel B of Figure 4 reports the coefficients of a regression of the probability score on lags and leads of implied skewness, along with 95% (pointwise) confidence bands. Prior to portfolio formation, higher implied skewness (i.e., optimistic expectations about asset returns) is associated with a lower probability score. However, the probability score is associated with higher future implied skewness. Put together, these findings corroborate our hypothesis that the D-HML portfolio profits from investors' biased expectations which are corrected upon information arrival.[12]

6 Additional Results and Robustness Checks

6.1 Behavioral Biases and Alternative Asset Pricing Models

We investigate whether the positive premium of the D-HML may be attributed to a set of well-known factors, including those capturing risk, mispricing, and behavioral biases. In particular, we estimate the abnormal returns (alphas) follow-

[12]We test whether an option-implied skewness factor built to capture the asymmetric expectations about future jumps in asset returns explains the return premium of the D-HML portfolio. Details on the factor construction can be found in Section D of the Online Appendix. Table A.7 of the Online Appendix shows that, once this factor is accounted for, the D-HML strategy does not command a positive and significant premium.

ing: 1) the Fama and French (1993) three-factor models (FF3); 2) the FF3 model augmented by the momentum, short-term reversal, and long-term reversal factors (FF3'); 3)the FF3' model augmented by the Pástor and Stambaugh (2003) liquidity factor (FF3"); 4) the Fama and French (2015) five-factor model augmented by the momentum, short-term reversal, long-term reversal, and liquidity factor (FF5'); 5)the Hou et al. (2015) four-factor model (Q4); 6) the Hou et al. (2020a) five-factor model augmented by the momentum, short-term reversal, long-term reversal, and liquidity factor (Q5'); 7) the Stambaugh and Yuan (2017) model (SY (2017)) with the MGMT and PERF mispricing factors; 8) the Daniel et al. (2020) model (DHS (2020)) with the FIN and PEAD behavioral factors; 9) the Asness et al. (2019) model (AFP (2019)) with the quality minus junk QMJ factor. We also show results obtained augmenting the FF5 model with: 10) the Bali et al. (2017) FMAX factor; 11) the Frazzini and Pedersen (2014) betting-against-beta (BAB) factor; 12) the Asness et al. (2000) betting-against-correlation (BAC) factor; 13) the betting-against-volatility (BAV) factor; 14) the Atilgan et al. (2020) left-tail momentum (LTM) factor; and 15) the idiosyncratic volatility IVOL factor estimated from the variance of residuals after the Fama and French (1993) three-factor model estimated on daily returns over the previous two months.

Models (1) to (15) in Table A.8 of the Online Appendix show the abnormal returns across asset pricing models. The positive premium associated with the D-HML portfolio remains significant after controlling for all the factors considered. Overall, our results suggest that the positive premium associated with the D-HML strategy is not fully explained by the existing mispricing factors, behavioral biases, lotteryness, and volatility factors.

6.2 Limits to Arbitrage

An alternative explanation for the D-HML premium is the existence of limits to arbitrage (Shleifer and Vishny (1997)). If there are limits to arbitrage, mispricing will tend to persist longer. We therefore expect the premium associated with D-HML to be larger among stocks with greater limits to arbitrage.

We test this hypothesis by calculating the abnormal return of the high PS minus low PS for subsamples of firms with low and high limits to arbitrage. We use 3 proxies to measure limits to arbitrage: i) firm size, measured as the market capitalization; ii) Amihud (2002) illiquidity; and iii) idiosyncratic volatility, calculated with the regression residuals obtained from the Fama and French (1993) three-factor model estimated with daily returns over the past two months. Arbitrage is more difficult for small and illiquid stocks with high idiosyncratic risk. Each month, firms are classified in high (bottom 33%) and low (top 33%) limits to arbitrage portfolios using each of the three proxies above. We then build the directional high-minus-low portfolios using the probability score for the high and low limits to arbitrage firms. Table A.9 in the Online Appendix presents the Fama and French (1993) three-factor (FF3) alphas for the 1932–2022 period and the Fama and French (2015) five-factor (FF5) alphas for the 1963–2022 period. The results reported in Table A.9 support the conjecture that the D-HML premium is more pronounced for stocks with high limits to arbitrage. The D-HML return premium is significantly larger for small, illiquid, and higher volatility firms. Nonetheless, the return premium is still economically and statistically significant for stocks with low limits to arbitrage, suggesting that limits to arbitrage provides only a partial explanation to the D-HML premium.

To investigate further whether limits to arbitrage in the form of short selling constraints can explain the D-HML return premium, we repeat the analysis carried out in Table 2 using exclusively firms with available traded options. Focusing exclusively on stocks with available options implies that we consider a subsample of stocks for which arbitrage is less restricted,[13] as investors could go long and

[13]As of December 2022, the Chicago Board Options Exchange (CBOE) imposes five criteria that a stock must meet before it can have options: (i) the underlying equity security must be a properly registered National Market System (NMS) stock, (ii) the company must have at least 7,000,000 publicly

short on the underlying equities or buy call and put options, should limits to short selling apply. Table A.10 in the Online Appendix contains the risk-adjusted returns of the D-HML and WML strategies. Abnormal returns from the D-HML remain significant at the 1% level.

6.3 Rank Accuracy

Self-financing strategies yield positive returns if the long portfolio has higher returns than the short portfolio. Past return characteristics, such as momentum (MOM) and the probability score (PS), are useful to the extent that they are able to accurately separate future underperformers from future outperformers. We verify the relevance of MOM and PS in predicting future return ranks, measured as percentiles, as well as in predicting future returns. In particular, we employ the Fama and MacBeth (1973) procedure and estimate the following regression:

$$z_{i,t+1} = \alpha_t + \sum_{\substack{d=1 \\ d \neq 5}}^{10} \beta_{d,t} \text{Decile}_{d,i,t} + u_{i,t} \ , \tag{7}$$

where $z_{i,t+1} \in [1, 100]$ represents the cross-sectional return percentile of stock i at time $t+1$ and $\text{Decile}_{d,i,t}$ is a dummy variable representing which decile (d) of the relevant characteristic the stock belongs to. We also estimate Eq. (7) using returns instead of percentiles as the dependent variable. Table A.11 in the Online Appendix reports the estimated coefficients. The constant term represents the 5^{th} decile.

Stocks in the 5^{th} decile of both PS and MOM are, on average, close to the median of the cross-sectional distribution of returns (constants of 50.03 and 50.93, respectively). Stocks in the bottom decile of PS are associated with a 3.06 percentile decrease in their rank, whereas being in the top decile of PS is associated with

held shares, (iii) the underlying stock must have at least 2,000 shareholders, (iv) trading volume must equal or exceed 2,400,000 shares in the past 12 months (v) the price of the security must be sufficiently high for a specific time. Altogether, these rules prevent small, volatile, and low liquid stocks from having options.

a 2.15 percentile increase in rank relative to stocks in the 5^{th} decile of PS. In contrast, stocks in the bottom decile of MOM are associated with a 4.68 percentile decrease in rank, whereas being in the top decile of MOM is associated with a statistically insignificant 0.14 percentile increase in rank relative to stocks in the 5^{th} decile of MOM. Overall, these results indicate that the probability score is useful in distinguishing future winners from losers.

We also report the future average realized return in each of the deciles of both PS and MOM. The right panel of Table A.11 highlights that stocks classified in the bottom PS decile have an average return decrease of -0.93%, whereas stocks in the top decile of PS have an average return increase of 0.78%, relative to those in the 5^{th} decile of PS. On the other hand, stocks classified in the bottom MOM decile have a statistically insignificant return decrease of -0.28%, whereas stocks in the top MOM decile have an average return increase of 0.48%, relative to those in the 5^{th} decile of MOM. Moving from the bottom to the top PS decile, stocks experience an average return increase of 1.71%, more than double relative to its MOM counterpart (0.76%). Overall, we find that the probability score is informative about future cross-sectional rankings as well as returns.

6.4 Robustness Checks

Transaction Costs We check whether the positive premium associated with the D-HML is the result of high transaction costs. To account for transaction costs, we calculate the returns of each strategy by using the bid and ask prices instead of the actual reported prices when initiating and closing positions. In particular, we assume that each position is opened at the mid-price between the bid and the ask, and closed with the worst price. In particular, when a long position is closed, the returns are computed as:

$$\widehat{r}_{i,t} = \frac{\mathrm{Bid}_{i,t} - \mathrm{Mid}_{i,t-1}}{\mathrm{Mid}_{i,t-1}} \ .$$

Analogously, when a short position is closed, returns are computed as:

$$\widehat{r}_{i,t} = \frac{\mathrm{Ask}_{i,t} - \mathrm{Mid}_{i,t-1}}{\mathrm{Mid}_{i,t-1}} \ ,$$

and portfolio returns are computed using $\widehat{r}_{i,t}$, the transaction cost adjusted returns. Table A.12 in the Online Appendix shows the raw and abnormal returns when transaction costs are accounted for. Although smaller, the abnormal returns associated with the D-HML strategy remain highly significant. Similar results are also observed for the WML strategy (Panel B of Table A.12), confirming the findings of Frazzini et al. (2014).

Small Stocks and Equally Weighted Portfolios To verify whether our results are driven by trading on small stocks, we repeat the analysis in Section 3 on the subset of stocks trading above \$5 per share at the portfolio formation month. Table A.13 in the Online Appendix reports our results. Both the D-HML and the WML strategies provide positive and significant risk-adjusted returns after removing small stocks. Results are qualitatively the same as those reported in Tables 2 and A.2.

We also test whether our results are sensitive to portfolio weighting schemes. Our results have so far been based on value-weighted portfolios. Table A.14 in the Online Appendix reports the results for equally weighted portfolios. The D-HML strategy reports higher abnormal returns, although with a substantially higher kurtosis and negative skewness. In contrast, the WML returns are substantially smaller, with a sharp increase in kurtosis and decrease in skewness.

Overall, the core findings presented in Section 3 remain consistent when small stocks are excluded or equally weighted portfolios are constructed.

Alternative Forecasting Models In addition to the model described in Eq. (1), we consider a number of alternative specifications. We estimate the probability of positive return using i) only duration variables, ii) duration variables and volatility, iii) duration variables and market return, iv) volatility and market return, v) the set of variables employed in Conrad et al. (2014) with and without the inclusion of our baseline variables, and vi) the set of firm characteristics employed in Kelly et al. (2021) with and without the inclusion of our baseline model's variables. Table A.15 in the Online Appendix contains the abnormal returns of the D-HML strategy built on alternative measures of the probability score (PS) estimated with alternative econometric specifications. The D-HML strategy produces positive and significant returns across a range of probability forecasting models considered, and the duration variables play a pivotal role in our results.

7 Conclusion

We document the existence of directional predictability in equity returns and introduce the directional high-minus-low (D-HML) strategy. The proposed strategy exploits directional information in the time series and cross-section of stock returns, and sorts stocks according to forecasts of the conditional probability of future positive returns, buying (selling) stocks with the highest (lowest) predicted probability of positive returns. Over roughly 100 years of U.S. equity market data, we find that the D-HML strategy generates a monthly abnormal return ranging from 0.76% to 1.16%, exhibits positively skewed returns with a small probability of large losses, and is not spanned by prominent factor models with the market, size, book-to-market, profitability, investment, momentum, and short/long-term reversal factors as well as mispricing and behavioral factors (Daniel et al. (2020); Stambaugh and Yuan (2017)), or the IPCA-based conditional asset pricing models (Kelly et al. (2021)). In addition, the Sharpe ratio of D-HML strategy is substantially larger than that of the momentum strategy (WML), with positively skewed returns that are also less prone to crashes, with its 0.1% quantile being equal to WML's 5% quantile.

We find empirical support to the claim that D-HML portfolio profits from mispricing due to investors' biased expectations (Engelberg et al. (2018)), with investors being excessively optimistic about some stocks and pessimistic about others, and we provide a simple theoretical model under which mispricing is generated from investors' biased expectations. Moreover, we find that the D-HML premium is significantly higher for i) stocks with higher presence of retail investors, ii) in months with earnings announcements, and iii) during high-sentiment periods in which investors are excessively optimistic.

References

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, 5:31–56.

Antoniou, C., Doukas, J. A., and Subrahmanyam, A. (2016). Investor sentiment, beta, and the cost of equity capital. *Management Science*, 62(2):347–367.

Asness, C., Friedman, J., Krail, R., and Liew, J. (2000). Style timing: Value versus growth. *Journal of Portfolio Management*, 26:50–60.

Asness, C. S., Frazzini, A., and Pedersen, L. H. (2019). Quality minus junk. *Review of Accounting Studies*, 24(1):34–112.

Atilgan, Y., Bali, T. G., Demirtas, K. O., and Gunaydin, A. D. (2020). Left-tail momentum: Underreaction to bad news, costly arbitrage and equity returns. *Journal of Financial Economics*, 135(3):725–753.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The Journal of Finance*, 61(4):1645–1680.

Bali, T., Brown, S., Murray, S., and Tang, Y. (2017). A lottery-demand-based explanation of the beta anomaly. *Journal of Financial and Quantitative Analysis*, 52:2369–2397.

Bali, T., Hu, J., and Murray, S. (2019). Option implied volatility, skewness, and kurtosis and the cross section of expected stock returns. *Working paper*.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343.

Barroso, P. and Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, 116(1):111–120.

Breen, W., Glosten, L. R., and Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *The Journal of Finance*, 44(5):1177–1189.

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

Conrad, J., Kapadia, N., and Xing, Y. (2014). Death and jackpot: Why do individual investors hold overpriced stocks? *Journal of Financial Economics*, 113(3):455–475.

Cumby, R. and Modest, D. M. (1997). Testing for market timing ability: A framework for forecast evaluation. *Journal of Financial Economics*, 19(1):35–55.

Daniel, K., Hirshleifer, D., and Sun, L. (2020). Short- and Long-Horizon Behavioral Factors. *Review of Financial Studies*, 33(4):1673–1736.

Daniel, K. and Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2):221–247.

Diebold, F. X. and Christoffersen, P. F. (2006). Financial asset returns, direction-of-change forecasting and volatility dynamics. *Management Science*, 52:1273–1287.

Diebold, F. X., Christoffersen, P. F., Mariano, R. S., Tay, A. S., and Tse, Y. K. (2007). Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence. *Journal of Financial Forecasting*, 1(2):1–22.

Diebold, F. X. and Lopez, J. A. (1996). Forecast evaluation and combination. *NBER*.

Engelberg, J., McLean, R. D., and Pontiff, J. (2018). Anomalies and news. *The Journal of Finance*, 73(5):1971–2001.

Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56.

Fama, E. and French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of finance*, 51:55–84.

Fama, E. and French, K. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116:1–22.

Fama, E. and MacBeth, J. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81:607–636.

Fama, E. F. (1991). Efficient capital markets: Ii. *The Journal of Finance*, 46(5):1575–1617.

Frazzini, A., Israel, R., and Moskowitz, T. (2014). Trading costs and asset pricing anomalies. *Chicago Booth Research Paper No. 14-05*.

Frazzini, A. and Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, 111:1–25.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting Characteristics Nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377.

Greene, W. H. (2017). *Econometric Analysis*. Pearson.

Grundy, B. D. and Martin, J. S. M. (2015). Understanding the Nature of the Risks and the Source of the Rewards to Momentum Investing. *The Review of Financial Studies*, 14(1):29–78.

Hou, K., Mo, H., Xue, C., and Zhang, L. (2020a). An Augmented q-Factor Model with Expected Growth*. *Review of Finance*, 25(1):1–41.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *Review of Financial Studies*, 28:650–705.

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

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2014). Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *The Review of Financial Studies*, 28(3):791–837.

Jegadeesh, N. and Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48:65–91.

Kelly, B. T., Moskowitz, T. J., and Pruitt, S. (2021). Understanding momentum and reversal. *Journal of Financial Economics*, 140(3):726–743.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Leung, M. T., Daouk, H., and Chen, A.-S. (2000). Forecasting stock indices: a comparison of classification and level estimation methods. *International Journal of Forecasting*, 16:173–190.

Linnainmaa, J. T. and Zhang, C. (2022). The earnings announcement return cycle. *SSRN*.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *Journal of Finance*, 20:587–615.

Maheu, J. M. and McCurdy, T. H. (2000). Identifying bull and bear markets in stock returns. *Journal of Business & Economics Statistics*, 18:100–112.

Malkiel, B. G. and Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work*. *The Journal of Finance*, 25(2):383–417.

Mixon, S. (2011). What does implied volatility skew measure? *Journal of Derivatives*, 18:1–9.

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica: Journal of the Econometric Society*, 34:768–783.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*, 78(2):277 – 309.

Newey, W. and West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708.

Pástor, L. and Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111:642–685.

Satchell, S. and Timmerman, A. (1995). An assessment of the economic value of non-linear foreign exchange rate forecasts. *Journal of Forecasting*, 14:477–497.

Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19:425–442.

Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1):35–55.

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.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302.

Stambaugh, R. F. and Yuan, Y. (2017). Mispricing Factors. *Review of Financial Studies*, 30(4):1270–1315.

Timmerman, A. and Lunde, A. (2004). Duration dependence in stock prices: An analysis of bull and bear markets. *Journal of Business & Economics Statistics*, 22:253–273.

Xing, Y., Zhang, X., and Zhao, R. (2010). What does the individual option volatility smirk tell us about future equity returns? *Journal of Financial and Quantitative Analysis*, 45(3):641 – 662.

Table 1:

Directional Predictability

This table reports coefficients obtained from the regression $y_{i,t+1} = a + bx_{i,t} + u_{i,t+1}$. The dependent variable $y_{i,t+1}$ is either a dummy variable for return directions ($r^+_{i,t+1}$) that takes the value of one if the return of stock i in month $t+1$ is strictly positive and zero otherwise, or the return ($r_{i,t+1}$) on the stock. In Models (1) and (3), $x_{i,t}$ is a binary dummy variable that is equal to 1 if the probability score ($\mathrm{PS}_{i,t}$) is higher than 0.5 ($\mathbb{1}\{\mathrm{PS}_{i,t} > 0.5\}$) and zero otherwise. In Models (2) and (4) the independent variable is the probability score ($\mathrm{PS}_{i,t}$). The first three rows report estimates obtained using Ordinary Least Squares (OLS), and the last three rows those obtained from Fama and MacBeth (1973) (FMB) regressions. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

		$r^+_{i\,t+1}$		$r_{i\,t+1}$	
		(1)	(2)	(3)	(4)
OLS	Constant	45.91***	0.05***	0.16***	−0.12***
		(547.09)	(8.57)	(7.24)	(−46.32)
	$\mathbb{1}\{\mathrm{PS}_t > 0.5\}$	4.51***		1.46***	
		(63.69)		(55.04)	
	PS_t		0.86***		0.26***
			(80.44)		(51.32)
FMB	Constant	45.98***	−0.02	0.57**	−0.10***
		(61.19)	(−0.65)	(2.05)	(−9.45)
	$\mathbb{1}\{\mathrm{PS}_t > 0.5\}$	5.23***		1.26***	
		(14.49)		(7.35)	
	PS_t		1.02***		0.23***
			(14.95)		(11.13)

Table 2:

Value-Weighted Portfolios sorted on PS

Panel A reports monthly raw and abnormal returns (alphas) of the directional high minus low portfolio (D-HML), obtained as the difference between the value-weighted high and low portfolios constructed on the probability score (PS). Abnormal returns (alphas) are generated based on the following models: (i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; (ii) the three-factor model of Fama and French (1993) with the MKT, SMB, and HML factors (FF3); (iii) the five-factor model of Fama and French (2015) with the MKT, SMB, HML, RMW, and CMA factors (FF5); (iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); (v) the Hou et al. (2020a) 5-factor model (Q5); (vi) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurtosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtosis of the portfolio returns. Panel B reports the FF3 abnormal returns for each value weighted portfolio built on deciles of PS. Max and Min indicate the maximum and minimum returns observed in each portfolio, respectively. Avg. Pos. Ret. is the average over time of the percentage of positive returns in each decile. Avg. Winners and Avg. Losers are the average over time of the percentage of stocks in each portfolio that are also in the top and bottom decile of future returns, respectively. Avg. Market Cap is the average over time of the median market capitalization (in thousands) of firms in each portfolio. Panel C reports the abnormal return for each decile portfolio using the asset pricing models considered in Panel A. The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5 model. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Panel A. Raw and abnormal returns (alphas) of the directional high minus low portfolios (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.77***	0.76***	0.82***	1.02***	0.92***	1.16***	0.93***
	(7.36)	(6.86)	(6.44)	(7.62)	(5.28)	(6.41)	(4.90)
MKT		0.01	0.03	0.05	−0.04	0.04	−0.03
		(0.19)	(0.95)	(1.24)	(−1.17)	(0.86)	(−0.82)
SMB			0.05	0.04	−0.04		
			(1.04)	(0.59)	(−0.71)		
HML			−0.19***	−0.05	−0.20**		
			(−3.37)	(−0.50)	(−2.03)		
RMW				−0.12*	−0.06		
				(−1.68)	(−0.74)		
CMA				−0.10	−0.05		
				(−0.84)	(−0.47)		
UMD					−0.07		−0.05
					(−1.39)		(−0.79)
STR					0.40***		0.39***
					(4.48)		(4.33)
LTR					0.16**		0.14**
					(2.19)		(2.11)
R_{ME}						0.02	−0.02
						(0.35)	(−0.38)
R_{IA}						−0.18**	−0.28***
						(−2.04)	(−2.68)
R_{ROE}						−0.10	0.00
						(−1.37)	(0.04)
R_{EG}						−0.07	0.04
						(−0.67)	(0.43)
T	1,086	1,086	1,086	714	714	672	672
R^2	0.05	0.05	0.09	0.10	0.25	0.11	0.24
Vol	3.24	3.24	3.24	3.32	3.32	3.38	3.38
SR	0.54	0.54	0.54	0.65	0.65	0.66	0.66
Skew	1.02	1.02	1.02	0.80	0.80	0.77	0.77
Kurtosis	10.52	10.52	10.52	5.86	5.86	5.71	5.71

Panel B. Abnormal returns for each decile portfolios built on PS

	1	2	3	4	5	6	7	8	9	10	10 − 1
Alpha	−0.43***	−0.29***	−0.13**	−0.09	−0.02	0.02	0.10**	0.10**	0.23***	0.39***	0.82***
	(-5.57)	(-4.98)	(-2.44)	(-1.56)	(-0.42)	(0.35)	(1.96)	(2.02)	(4.47)	(7.24)	(6.44)
MKT	0.99***	1.00***	1.04***	1.01***	1.02***	1.04***	0.99***	1.03***	0.95***	1.01***	0.03
	(44.47)	(64.42)	(41.40)	(65.29)	(74.70)	(51.41)	(55.86)	(56.60)	(36.86)	(58.92)	(0.95)
SMB	0.02	0.00	0.03	0.00	0.02	0.04	0.04	0.11**	0.03	0.07*	0.05
	(0.36)	(0.18)	(1.02)	(0.15)	(0.71)	(1.48)	(1.06)	(2.23)	(1.29)	(1.86)	(1.04)
HML	0.14***	0.08***	0.14***	0.08***	0.06***	0.02	0.02	0.11***	−0.03	−0.05**	−0.19***
	(3.42)	(3.40)	(4.09)	(3.71)	(3.16)	(0.67)	(0.51)	(4.48)	(-1.10)	(-2.08)	(-3.37)
Vol	5.59	5.42	5.74	5.48	5.55	5.69	5.42	5.74	5.16	5.53	3.24
SR	0.25	0.34	0.46	0.47	0.51	0.52	0.57	0.59	0.65	0.74	0.54
Skew	1.21	0.69	1.64	0.51	0.99	0.76	0.66	1.43	0.19	0.4	1.02
Max	0.51	0.46	0.58	0.41	0.49	0.46	0.47	0.45	0.39	0.37	0.28
Min	-0.3	-0.3	-0.27	-0.27	-0.25	-0.29	-0.24	-0.26	-0.22	-0.23	-0.17
Avg. Pos. Ret.	0.46	0.48	0.48	0.49	0.5	0.5	0.51	0.52	0.53	0.53	–
Avg. Winners	0.10	0.09	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.11	–
Avg. Losers	0.13	0.11	0.10	0.09	0.09	0.09	0.09	0.08	0.08	0.09	–
Avg. Market Cap	189,220	182,391	178,853	170,654	170,425	185,788	186,120	196,995	194,804	167,844	–

Panel C. Abnormal Returns for each decile portfolios built on PS (alternative asset pricing models)

	1	2	3	4	5	6	7	8	9	10	10 − 1
CAPM	−0.39***	−0.26***	−0.09*	−0.07	−0.00	0.03	0.10*	0.14***	0.22***	0.38***	0.76***
	(-5.39)	(-4.81)	(-1.75)	(-1.17)	(-0.07)	(0.50)	(1.95)	(2.61)	(4.26)	(6.55)	(6.86)
FF5	−0.50***	−0.33***	−0.17**	−0.13*	−0.06	0.07	0.15**	0.16**	0.25***	0.52***	1.02***
	(-5.83)	(-4.43)	(-2.27)	(-1.95)	(-1.01)	(1.11)	(2.10)	(2.42)	(3.71)	(6.97)	(7.62)
FF5+UMD+S(L)TR	−0.44***	−0.30***	−0.11	−0.14*	−0.06	0.08	0.17**	0.14*	0.23***	0.48***	0.92***
	(-4.00)	(-3.88)	(-1.17)	(-1.89)	(-0.99)	(1.07)	(2.17)	(1.76)	(3.28)	(5.19)	(5.28)
Q5	−0.51***	−0.38***	−0.13	−0.04	−0.07	0.14*	0.15**	0.16**	0.22**	0.65***	1.16***
	(-4.75)	(-4.70)	(-1.33)	(-0.58)	(-0.98)	(1.74)	(2.01)	(2.07)	(2.57)	(6.44)	(6.41)
Q5+UMD+S(L)TR	−0.42***	−0.33***	−0.08	−0.06	−0.07	0.15	0.12	0.13	0.18**	0.51***	0.93***
	(-3.40)	(-3.64)	(-0.76)	(-0.81)	(-0.99)	(1.63)	(1.40)	(1.60)	(2.25)	(5.30)	(4.90)

Table 3:

Risk and the PS Portfolios

This table reports the averages of i) the market beta (Market Beta), ii) the market beta with respect to negative and positive market returns (Neg. iii) Beta and Pos. Beta), iv) the total volatility of returns (Total Vol.), v) the downside volatility (Neg. Vol.), vi) the upside volatility (Pos. Vol.), vii) the idiosyncratic volatility (Idio. Vol.) computed from CAPM residuals, viii) the downside volatility of idiosyncratic returns (Neg. Idio. Vol.) and ix) the upside volatility of idiosyncratic returns (Pos. Idio. Vol.) for stocks in the top (High PS) and bottom (Low PS) deciles of the probability score. Portfolios are formed at month t and kept until month $t+1$. The pre-formation period corresponds to the two-month period from $t-1$ to t. The holding period covers from month t to month $t+1$, and the post-formation period runs from month $t+1$ to $t+2$. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	Pre Formation			Holding Period			Post Formation		
	Low PS	High PS	Diff.	Low PS	High PS	Diff.	Low PS	High PS	Diff.
Market Beta	0.83***	0.80	−0.03***	0.84***	0.81	−0.04***	0.82***	0.83	0.01
	(145.41)		(−3.89)	(145.69)		(−5.01)	(143.90)		(1.33)
Neg. Beta	0.97***	0.86	−0.10***	0.98***	0.86	−0.12***	0.92***	0.91	−0.01
	(126.59)		(−11.15)	(123.33)		(−12.40)	(126.72)		(−1.21)
Pos. Beta	0.71***	0.74	0.02**	0.71***	0.75	0.04***	0.73***	0.76	0.03***
	(87.86)		(2.28)	(84.90)		(3.53)	(94.67)		(3.31)
Total Vol.	0.05***	0.03	−0.02***	0.05***	0.03	−0.01***	0.04***	0.03	−0.01***
	(86.11)		(−27.20)	(85.91)		(−24.84)	(91.17)		(−13.67)
Neg. Vol.	0.03***	0.02	−0.01***	0.03***	0.02	−0.01***	0.02***	0.02	−0.00***
	(93.23)		(−20.17)	(93.91)		(−19.28)	(100.52)		(−13.63)
Pos. Vol.	0.04***	0.02	−0.02***	0.04***	0.02	−0.02***	0.03***	0.03	−0.01***
	(79.48)		(−32.95)	(79.91)		(−29.00)	(84.40)		(−12.59)
Idio. Vol.	0.04***	0.03	−0.02***	0.04***	0.03	−0.01***	0.04***	0.03	−0.01***
	(82.25)		(−26.97)	(82.21)		(−24.68)	(86.38)		(−13.67)
Neg. Idio. Vol.	0.02***	0.02	−0.01***	0.02***	0.02	−0.01***	0.02***	0.02	−0.00***
	(87.89)		(−20.69)	(88.80)		(−19.37)	(94.06)		(−13.43)
Pos. Idio. Vol.	0.04***	0.02	−0.02***	0.04***	0.02	−0.02***	0.03***	0.02	−0.01***
	(76.10)		(−33.44)	(76.53)		(−30.02)	(80.11)		(−13.67)

Table 4:

D-HML Returns and Conditional Risk Models

Panel A presents the coefficient estimates of multivariate panel regressions of next month's excess stock returns $r_{i,t+1}$ on the current month's signal (Models (1) to (4)) or signal rank (Models (5) to (8)) built on MOM, IPCA and PS. t-statistics are calculated using standard errors clustered by month. The R-square values are multiplied by 100 for better interpretation. Slope coefficients in rank-based regressions are also multiplied by 100 to enhance readability. Panel B reports the directional high minus low portfolio (D-HML) constructed using the next month residual returns calculated as the difference between the actual and the IPCA return for each firm i, i.e., $\epsilon_{i\,t+1} = r_{i\,t+1} - IPCA_{i\,t}$. Abnormal returns (alphas) are generated based on the following models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with the MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with the MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); v) the Hou et al. (2020a) 5-factor model (Q5); vi) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurtosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtosis of the portfolio returns. The sample covers the period from 1964 to 2022. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Panel A. Multivariate panel regressions with dependent variable next month return $r_{i,t+1}$

	Signal				Signal Rank			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.00	−0.12***	−0.11***	−0.11***	0.01***	0.01***	0.01***	0.01***
	(1.08)	(−3.66)	(−3.49)	(−3.46)	(4.38)	(4.36)	(4.36)	(4.35)
MOM	−0.01	−0.00		−0.00	−0.14	0.61**		−0.10
	(−1.39)	(−0.21)		(−1.34)	(−0.53)	(2.47)		(−0.36)
IPCA	0.86***		0.79***	0.82***	4.07***		3.86***	3.88***
	(12.28)		(11.77)	(11.14)	(12.73)		(12.72)	(12.03)
PS		0.25***	0.22***	0.22***		1.95***	1.48***	1.48***
		(4.05)	(3.61)	(3.59)		(17.10)	(13.11)	(13.05)
R^2 (x100)	0.64	0.32	0.88	0.90	0.61	0.16	0.69	0.69

Panel B. Raw and abnormal residual returns (alphas) of the directional high minus low portfolios (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.31** (2.47)	0.32** (2.57)	0.34*** (2.77)	0.38*** (2.93)	0.32** (2.38)	0.38*** (2.81)	0.29** (2.07)
MKT		−0.02 (−0.68)	−0.04 (−1.19)	−0.04 (−1.31)	−0.07** (−2.34)	−0.04 (−1.25)	−0.07** (−2.05)
SMB			0.03 (0.57)	−0.00 (−0.03)	−0.03 (−0.63)		
HML			−0.07 (−1.45)	−0.06 (−0.93)	−0.11 (−1.48)		
RMW				−0.09 (−1.23)	−0.07 (−0.96)		
CMA				−0.01 (−0.13)	−0.00 (−0.01)		
UMD					0.00 (0.13)		0.02 (0.50)
STR					0.16*** (2.86)		0.16*** (2.79)
LTR					0.06 (0.91)		0.04 (0.72)
R_{ME}						0.02 (0.36)	−0.01 (−0.09)
R_{IA}						−0.09 (−1.28)	−0.11 (−1.63)
R_{ROE}						−0.02 (−0.31)	−0.01 (−0.10)
R_{EG}						−0.01 (−0.13)	0.02 (0.28)
T	641	641	641	641	641	641	641
R^2	0.01	0.01	0.01	0.01	0.03	0.01	0.03
Vol	3.04	3.04	3.04	3.04	3.04	3.04	3.04
SR	0.35	0.35	0.35	0.35	0.35	0.35	0.35
Skew	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
Kurtosis	4.97	4.97	4.97	4.97	4.97	4.97	4.97

Table 5:

Probability Score and Earnings Announcements

The table reports results from a panel regression of next month excess return on the probability score (PS) and its interaction with a binary dummy taking value 1 if in a given month there has been an earnings announcement (Earnings Month), and zero otherwise. For comparison, Models (2)-(3) and (5)-(6) include the *Net* measure calculated as in Engelberg et al. (2018) using a set of 171 market anomalies (Chen and Zimmermann (2022)). Controls include 3 months lags of volatility and excess returns, as well as month fixed effects. N is the total number of firm-month observations and t-statistics are reported in parenthesis. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

	(1)	(2)	(3)	(4)	(5)	(6)
Earnings Month	0.24**	0.22**	0.24**	0.24**	0.21**	0.24**
	(2.28)	(2.11)	(2.25)	(2.25)	(2.03)	(2.24)
PS	0.68***		0.68***	0.48***		0.48***
	(6.09)		(6.10)	(5.00)		(5.02)
PS $\times$ Earnings Month	0.23**		0.24**	0.20**		0.21**
	(2.21)		(2.31)	(2.05)		(2.14)
Net		0.64***	0.63***		0.65***	0.64***
		(9.12)	(9.17)		(9.79)	(9.83)
Net $\times$ Earnings Month		0.24***	0.24***		0.24***	0.24***
		(3.16)	(3.25)		(3.12)	(3.21)
Controls	No	No	No	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
N	2,224,953	2,374,025	2,224,953	2,207,392	2,314,675	2,207,392

Table 6:

Mispricing: *Net* Based

The table reports the abnormal returns (alphas) of the low, high, and high minus low portfolios based on terciles of the *Net*, the PS, and the intersection of both measures. Abnormal returns (alphas) are generated based on the following models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the five-factor model of Fama and French (2015) with the MKT, SMB, HML, RMW, and CMA factors (FF5); iii) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); iv) the Hou et al. (2020a) 5-factor model (Q5); v) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). The *Net* measure is calculated as in Engelberg et al. (2018) using a set of 171 market anomalies (Chen and Zimmermann (2022)). The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5 model. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	Net			PS			PS ∩ *Net*		
	Low	High	High-Low	Low	High	High-Low	Low	High	High-Low
CAPM	-0.23^{***}	0.48^{***}	0.71^{***}	-0.28^{***}	0.16^{***}	0.44^{***}	-0.45^{***}	0.76^{***}	1.21^{***}
	(-6.58)	(8.07)	(8.90)	(-6.06)	(3.91)	(5.74)	(-8.42)	(10.77)	(11.02)
FF5	-0.11^{***}	0.26^{***}	0.37^{***}	-0.32^{***}	0.25^{***}	0.58^{***}	-0.44^{***}	0.56^{***}	1.00^{***}
	(-3.37)	(5.89)	(5.92)	(-4.96)	(4.83)	(6.32)	(-5.89)	(8.25)	(8.22)
FF5+UMD+S(L)TR	-0.07^{*}	0.21^{***}	0.28^{***}	-0.26^{***}	0.26^{***}	0.51^{***}	-0.36^{***}	0.50^{***}	0.86^{***}
	(-1.73)	(4.25)	(4.08)	(-3.15)	(4.56)	(3.92)	(-3.90)	(6.58)	(6.08)
Q5	-0.07	0.19^{***}	0.26^{***}	-0.28^{***}	0.32^{***}	0.60^{***}	-0.40^{***}	0.47^{***}	0.86^{***}
	(-1.57)	(2.89)	(2.79)	(-3.98)	(4.99)	(5.28)	(-4.56)	(4.92)	(6.04)
Q5+UMD+S(L)TR	-0.08	0.18^{***}	0.26^{**}	-0.22^{***}	0.25^{***}	0.47^{***}	-0.34^{***}	0.43^{***}	0.76^{***}
	(-1.56)	(2.62)	(2.55)	(-2.66)	(4.17)	(3.86)	(-3.32)	(4.53)	(4.78)

Table 7:

Mispricing: IPCA

Panel A of this table reports the average IPCA residual across PS deciles. IPCA residuals are constructed as $\epsilon_{i,t} = r_{i,t} - \widehat{r}_{i,t}$, where $r_{i,t}$ is the excess returns on stock i at time t, and $\widehat{r}_{i,t}$ is the IPCA prediction for the excess returns of stock i at time t. IPCA predictions are constructed out-of-sample. Panel B reports the abnormal returns for 25 value-weighted portfolios obtained by double sorting stocks using lagged IPCA residuals and the PS. IPCA residuals are calculated using out-of-sample IPCA predictions. Abnormal returns are calculated using the Fama and French (1993) 3-factor model (FF3). The sample period covers July 1969 to November 2022. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Panel A. Average IPCA residuals across PS deciles

	Low PS	2	3	4	5	6	7	8	9	High PS	High−Low (PS)
Avg. Residual	−0.69***	−0.42***	−0.29***	−0.07*	0.01	0.11***	0.23***	0.30***	0.40***	0.62***	1.31***
	(-11.38)	(-8.62)	(-5.62)	(-1.74)	(0.40)	(2.65)	(5.71)	(6.28)	(8.13)	(14.21)	(15.03)

Panel B. Double sorting portfolios on lag IPCA residuals and PS.

	Low PS	2	3	4	High PS	High-Low (PS)
Low Residuals	−0.91***	−0.34**	0.04	0.26*	0.57***	1.49***
	(−4.52)	(−2.19)	(0.27)	(1.72)	(3.48)	(6.16)
2	−0.20	0.16	0.13	0.31**	0.52***	0.72***
	(−1.25)	(1.25)	(1.19)	(2.51)	(4.70)	(3.83)
3	0.11	−0.05	0.10	0.23***	0.36***	0.25*
	(1.25)	(−0.50)	(1.09)	(2.71)	(3.70)	(1.79)
4	−0.37***	−0.18*	−0.09	0.04	0.15	0.52***
	(−3.77)	(−1.83)	(−0.81)	(0.39)	(1.29)	(3.28)
High Residuals	−0.70***	−0.59***	−0.23**	−0.20*	0.41**	1.14***
	(−5.67)	(−4.89)	(−2.00)	(−1.76)	(2.19)	(4.45)
High - Low (Residuals)	0.21	−0.25	−0.26	−0.46**	−0.19	
	(0.96)	(−1.23)	(−1.42)	(−2.28)	(−0.77)	

Table 8:

Probability Score and Institutional Ownership

The table reports in Panel A the abnormal returns (alphas) of 25 value-weighted portfolios obtained from independent double sorting firms based on the percentage of institutional ownership (%IO) and their probability score PS. Every month we independently double sort stocks into five equally spaced quintiles based on the percentage of institutional ownership observed at the end of the previous quarter and on our probability score PS. Then, we form 25 value weighted portfolios from the intersection of the two measures. Abnormal returns are obtained based on the Fama and French (2015) 5-factor model augmented with momentum (UMD), short-term reversal (STR) and long-term reversal (LTR). Panel B reports the abnormal returns (alphas) of the low-PS, high-PS and the directional high minus low signal portfolios (D-HML) obtained as difference between high-PS and low-PS for the high (top 33%) and low (bottom 33%) institutional ownership. Institutional ownership is measured using the i) total reported institutional ownership over the last quarter (IO), ii) the percentage of the institutional ownership (%IO) and iii) the residual institutional ownership estimated as in Nagel (2005) (R-IO). Abnormal returns are calculated using the Fama and French (1993) 3-factor model (FF3) and the Fama and French (2015) 5-factor model (FF5). The sample period is from from January 1980 to June 2022. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

Panel A. Abnormal returns of 25 double sorted portfolios

	Low PS	2	3	4	High PS	High-Low (PS)
Low (% Own.)	-0.70^{***}	-0.09	0.19	0.39^{**}	0.53^{***}	1.23^{***}
	(-3.37)	(-0.65)	(1.22)	(2.53)	(2.72)	(4.62)
2	-0.18	0.11	0.61^{***}	0.33^{**}	0.82^{***}	0.99^{***}
	(-0.99)	(0.84)	(4.01)	(2.48)	(4.30)	(3.37)
3	-0.18	0.32^{**}	0.35^{***}	0.54^{***}	0.87^{***}	1.05^{***}
	(-1.28)	(2.31)	(2.64)	(3.84)	(5.27)	(4.56)
4	0.04	0.21^{**}	0.24^{***}	0.58^{***}	0.72^{***}	0.68^{***}
	(0.43)	(2.11)	(2.78)	(6.41)	(6.15)	(4.09)
High (% Own.)	-0.12	0.08	0.18^{*}	0.46^{***}	0.51^{***}	0.63^{***}
	(-1.08)	(0.83)	(1.81)	(4.16)	(4.23)	(3.77)
High - Low (% Own.)	0.58^{***}	0.18	-0.01	0.07	-0.02	
	(2.67)	(0.99)	(-0.04)	(0.36)	(-0.09)	

Panel B. Differential abnormal returns for high and low institutional ownership (regressions)

		Low PS			High PS			D-HML		
		High IO	Low IO	Diff.	High IO	Low IO	Diff.	High IO	Low IO.	Diff.
FF3	IO	-0.49^{***}	-0.94	-0.45^{**}	0.34^{***}	1.06	0.72^{***}	0.83^{***}	2.00	1.17^{***}
		(-5.08)		(-2.15)	(3.06)		(3.68)	(6.59)		(5.28)
	%IO	-0.62^{***}	-1.31	-0.69^{***}	0.40^{***}	1.07	0.67^{***}	1.02^{***}	2.37	1.35^{***}
		(-6.78)		(-3.11)	(3.47)		(2.78)	(8.27)		(5.43)
	R-IO	-0.82^{***}	-1.24	-0.42^{**}	0.77^{***}	0.56	-0.21	1.59^{***}	1.80	0.21
		(-6.29)		(-2.18)	(5.47)		(-1.14)	(9.34)		(0.87)
FF5	IO	-0.53^{***}	-0.68	-0.16	0.33^{***}	1.15	0.82^{***}	0.86^{***}	1.84	0.98^{***}
		(-5.22)		(-0.65)	(3.17)		(3.84)	(6.59)		(4.19)
	%IO	-0.63^{***}	-0.95	-0.32	0.37^{***}	1.31	0.94^{***}	0.99^{***}	2.25	1.26^{***}
		(-6.47)		(-1.21)	(3.53)		(3.37)	(7.61)		(4.53)
	R-IO	-0.75^{***}	-1.00	-0.25	0.76^{***}	0.71	-0.05	1.52^{***}	1.71	0.19
		(-5.41)		(-1.09)	(4.81)		(-0.23)	(8.71)		(0.78)

Table 9:

Portfolio Analysis and Investor Sentiment

The table reports the abnormal returns (alphas) of the low-PS, high-PS and the directional high minus low signal portfolios (D-HML) obtained as difference between high-PS and low-PS. Abnormal returns are computed based on the Fama and French (2015) 5-factor model augmented with momentum (UMD), short-term reversal (STR) and long-term reversal (LTR). The variable low-sentiment (high-sentiment) is a binary dummy taking value 1 if the previous month the sentiment index is below (above) the median point and zero otherwise. We use three alternative sentiment indices: i) the Baker and Wurgler (2006) index, ii) the University of Michigan Consumer Sentiment Index and iii) the Huang et al. (2014) news based sentiment index. Newey-West adjusted t-statistics are given in parentheses. The sample period is from from July 1965 to June 2022 for the Baker and Wurgler (2006), January 1978 to April 2022 for the University of Michigan and August 1965 to December 2020 for the Huang et al. (2014) sentiment index, respectively. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

	Baker and Wurgler (2006)			University of Michigan			Huang et al. (2014)		
	Low-PS	High-PS	D-HML	Low-PS	High-PS	D-HML	Low-PS	High-PS	D-HML
Low-Sentiment	−0.21	0.41***	0.62***	−0.28*	0.46***	0.74***	−0.24**	0.34***	0.59***
	(−1.60)	(4.04)	(3.23)	(−1.78)	(4.43)	(3.58)	(−2.15)	(3.50)	(3.28)
High-Sentiment	−0.71***	0.63***	1.33***	−0.77***	0.63***	1.39***	−0.65***	0.66***	1.31***
	(−5.36)	(5.82)	(6.15)	(−5.53)	(4.57)	(5.13)	(−4.10)	(5.67)	(5.26)
MKT	1.03***	0.98***	−0.05	1.03***	0.98***	−0.06	1.02***	0.98***	−0.04
	(41.10)	(48.29)	(−1.45)	(36.15)	(42.98)	(−1.35)	(39.92)	(46.53)	(−1.16)
SMB	0.05	0.01	−0.03	0.07*	−0.03	−0.10*	0.05	0.01	−0.04
	(1.38)	(0.41)	(−0.67)	(1.80)	(−0.86)	(−1.72)	(1.53)	(0.29)	(−0.82)
HML	0.11*	−0.10*	−0.21**	0.10	−0.14**	−0.24**	0.13**	−0.11*	−0.24**
	(1.92)	(−1.84)	(−2.04)	(1.62)	(−2.44)	(−1.99)	(2.38)	(−1.91)	(−2.35)
RMW	0.09	0.00	−0.09	0.08	0.00	−0.08	0.08	0.00	−0.08
	(1.34)	(0.02)	(−1.01)	(1.19)	(0.02)	(−0.94)	(1.21)	(0.02)	(−0.91)
CMA	0.09	0.01	−0.08	0.13*	0.04	−0.09	0.05	0.02	−0.03
	(1.26)	(0.18)	(−0.63)	(1.70)	(0.55)	(−0.66)	(0.67)	(0.19)	(−0.25)
UMD	0.00	−0.07**	−0.07	−0.02	−0.08**	−0.06	−0.01	−0.07*	−0.06
	(−0.15)	(−2.04)	(−1.38)	(−0.59)	(−2.01)	(−1.15)	(−0.42)	(−1.91)	(−1.15)
STR	−0.16***	0.24***	0.40***	−0.14***	0.22***	0.35***	−0.17***	0.25***	0.41***
	(−3.62)	(4.88)	(4.83)	(−2.97)	(4.03)	(3.83)	(−3.49)	(4.87)	(4.53)
LTR	−0.15***	0.01	0.16**	−0.20***	0.02	0.21**	−0.14***	0.00	0.14*
	(−3.13)	(0.14)	(2.10)	(−3.47)	(0.39)	(2.54)	(−2.78)	(−0.06)	(1.75)
T	683	683	683	531	531	531	665	665	665
R^2	0.84	0.88	0.27	0.83	0.83	0.83	0.84	0.89	0.28

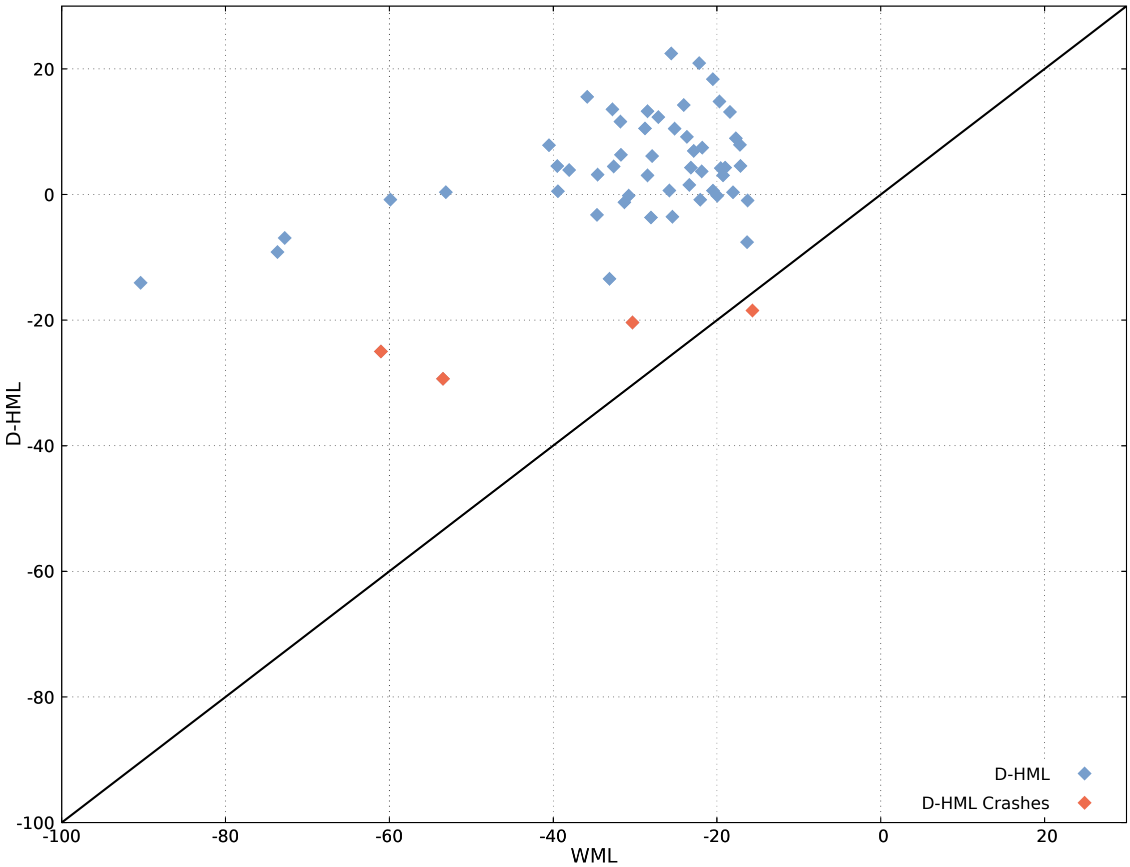


Figure 1:

Crashes and Cumulative Returns

This figure contains the scatter of the 3-month cumulative returns of the D-HML (on the vertical axis) against the WML strategy (on the horizontal axis) conditional on being in a momentum crash (either strategy displaying 3-month cumulative returns below the 5th percentile of the WML). Blue (orange) marks indicate that the WML (D-HML) strategy suffered a crash. Crashes are defined as episodes where the 3-month cumulative excess return of either strategy lies below momentum 5% quantile (-16.20%).

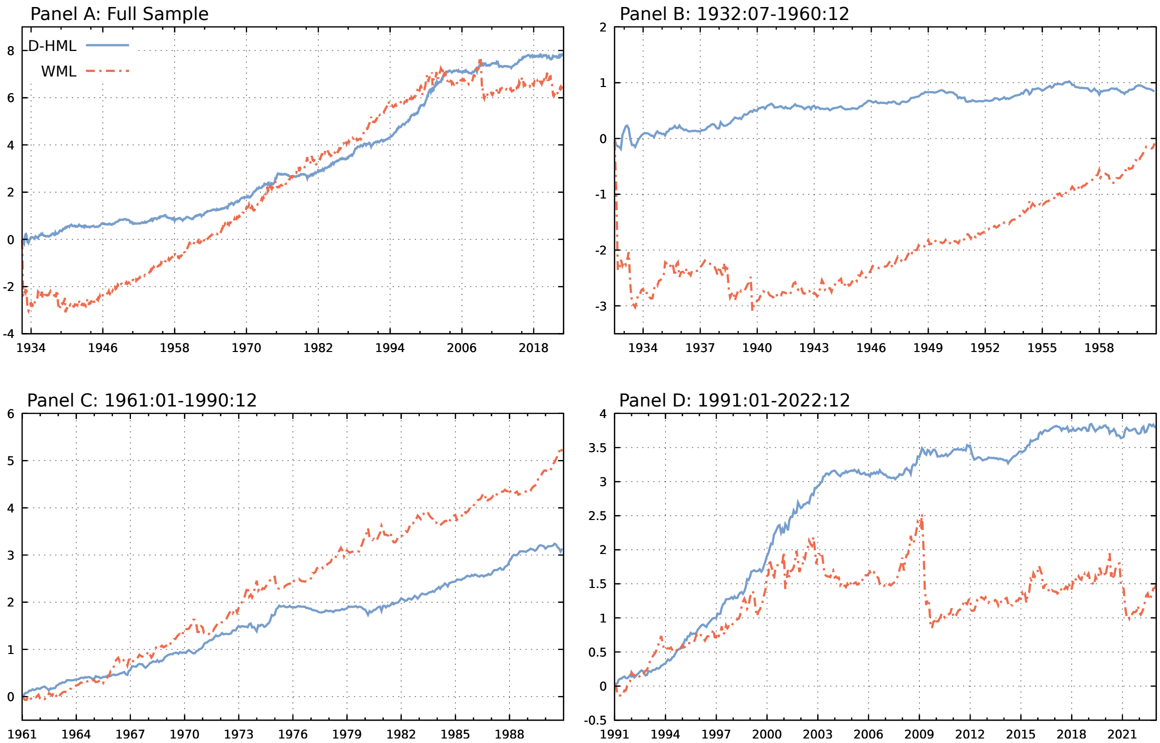
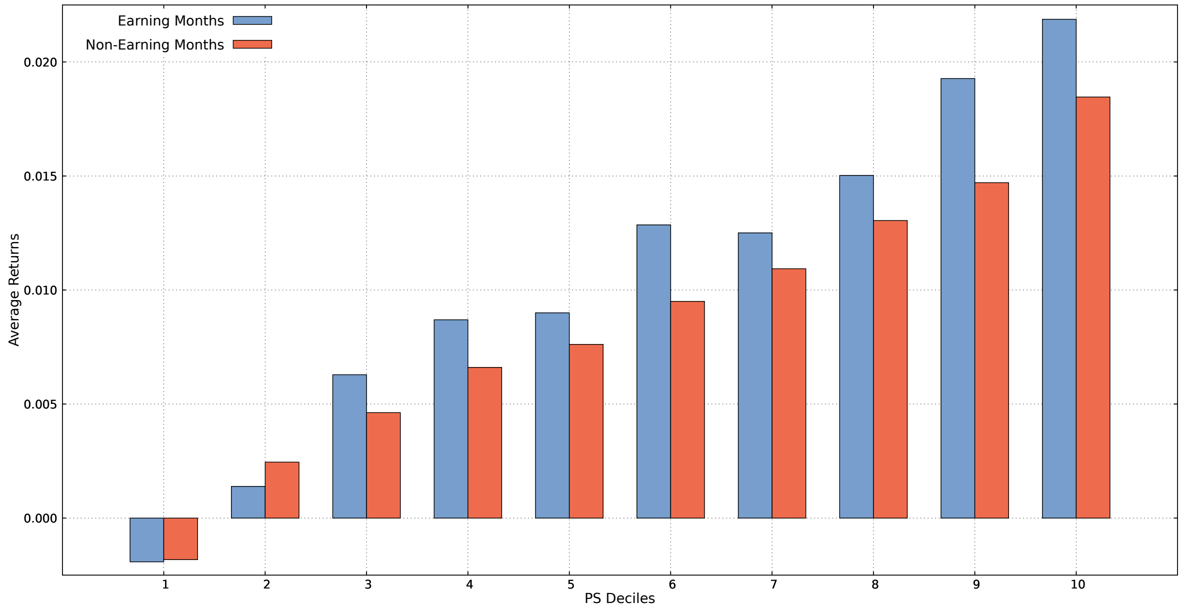


Figure 2:

Long-Term Performance

This figure contains the cumulative log return of the D-HML (blue continuous line) and the WML (orange dashed line). Panel A plots the cumulative returns over the full sample period from 1932:07 to 2022:12. Panels B–D plot the returns over three roughly 30 years sub-samples: 1932–1960, 1961–1990 and 1991–2022.

Panel A. Returns and earnings announcements



Panel B. Returns and earnings announcements for high vs. low sentiment periods

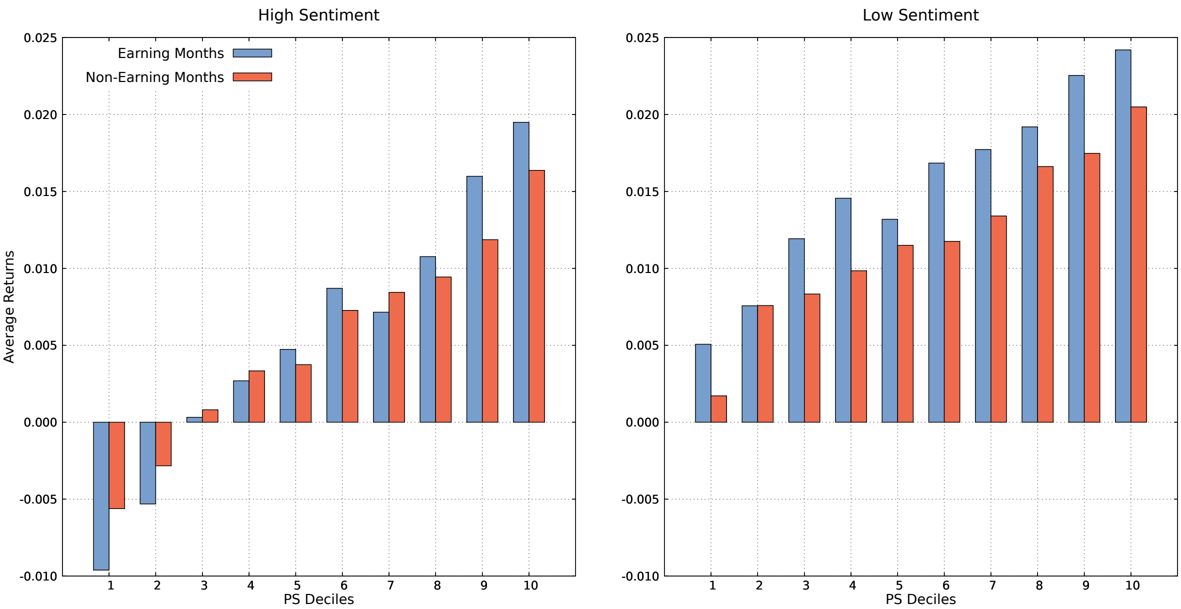


Figure 3:

Earnings and Returns

This figure contains in Panel A the average return of stocks that have earnings announcements and those that do not, grouped by deciles of Probability Score. Panel B contains the average returns of stocks that have earnings announcements for high (left panel) vs. low (right panel) sentiment periods measured using the Baker and Wurgler (2006) investor sentiment index.

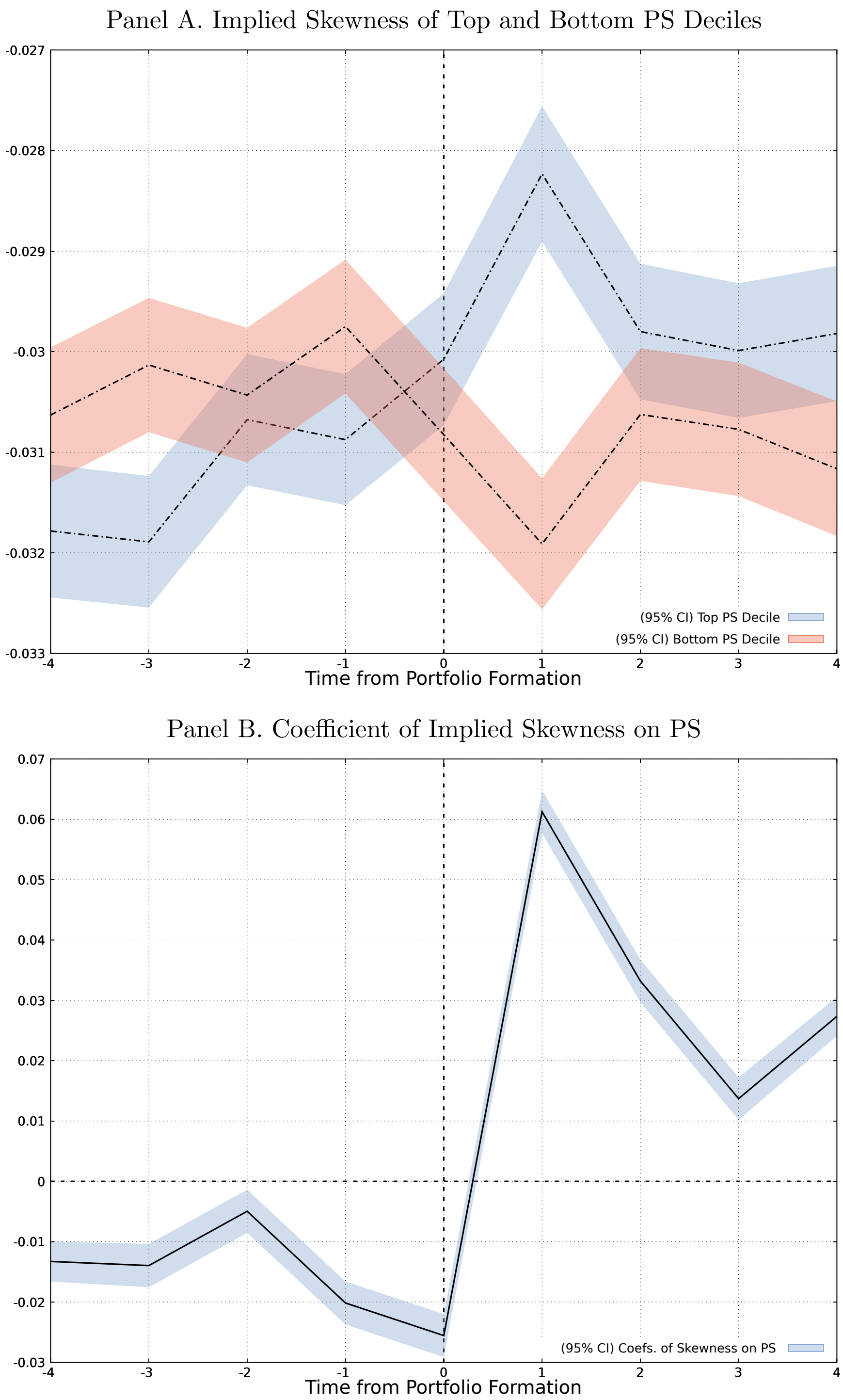


Figure 4:

PS and Implied Skewness

Panel A contains the average implied skewness of stocks in the top vs. the bottom deciles of the probability forecasts in a 4 month window around portfolio formation. Panel B contains the estimated coefficients of a regression of the probability forecasts (PS) on implied skewness.

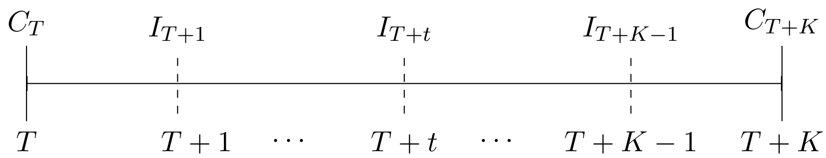
Appendix

Directional Information in Equity Returns

Luca Del Viva, Carlo Sala and André B.M. Souza

A A Model of Investors Biased Expectations

In this section we build on Barberis et al. (1998) and develop a theoretical model for investors' biased expectations. Suppose that the earnings at time $T + K$ are given by $C_{T+K} = C_T + z_{T+K}$ where $z_{T+K} \sim gN(\mu_G, 1) + (1 - g)N(\mu_B, 1)$ where $g \sim Bernoulli(p_G)$ and $\mu_G > \mu_B$. Assume all earnings are paid out as dividends. Between periods T and $T + K$ there are K sub-periods t, and the investor observes a sequence of informational shocks I_t, which we model as $I_t = 2h_t - 1$, where $h_t \sim Bernoulli(p_I)$. There are two possible states of the world: i) bad, where firm earnings are distributed as $N(\mu_B, 1)$ and ii) good, where earnings are distributed as $N(\mu_G, 1)$. The probability of being in the good state is p_G. Crucially, p_G is unknown to the investor and does not depend on $\{I_t\}_{t=T}^{T+K}$, but $\{I_t\}_{t=T}^{T+K}$ may depend on p_G through p_I. The investor does not know the probabilities of each state (p_G), and attempts to learn such probabilities based on the sequence of news he receives. We assume expectations are built according to a Markov process where good (bad) news reinforce his beliefs that he is in the good (bad) state.



Given two information shocks, the investor updates his beliefs at each point in time according to the following probabilities: where $0 < \gamma < 1$, $0 < \theta < 1$ and

Good	$I_{t+1} = 1$	$I_{t+1} = -1$
$I_t = 1$	π_1	$1 - \pi_1$
$I_t = -1$	$1 - \theta\pi_2$	$\theta\pi_2$

Bad	$I_{t+1} = 1$	$I_{t+1} = -1$
$I_t = 1$	$\gamma\pi_1$	$1 - \gamma\pi_1$
$I_t = -1$	$1 - \pi_2$	π_2

	$s_{t+1} = \text{Good}$	$s_{t+1} = \text{Bad}$
$s_t = \text{Good}$	$1 - \lambda$	λ
$s_t = \text{Bad}$	λ	$1 - \lambda$

$\pi_1, \pi_2, \lambda \in [0,1]$. As in Barberis et al. (1998) we assume that the investor knows π_1, π_2, θ, γ and λ. In words, the investor understands that information shocks are random, but believes that the likelihood of observing positive news depends on the underlying state, which he does not know and must learn in order to correctly price the asset. In particular, the investor believes that a sequence of consecutive positive shocks happens with probability π_1 in the good state, and with probability $\gamma\pi_1 < \pi_1$ in the bad state. Conversely, the investor attributes probability $\theta\pi_2$ of observing two consecutive negative shocks in the good state, and $\pi_2 > \theta\pi_2$ in the bad state. At period $T+K$ the earnings are paid out according to the state. Therefore, just before the earnings announcement, the investors problem is to identify whether he is in the good or bad state. Let $q_t = \mathbb{P}(s_t = \text{Good}|I_t, I_{t-1}, q_{t-1})$. For simplicity, we assume the interest rate between periods T and $T+K$ to be 0. If we assume the investor builds expectations following a Markov process as in Barberis et al. (1998), the stock price implied by the investor expectations at $T+K-1$ will be.

$$\begin{aligned}
P_{T+K-1|\{I_t\}_{t=T}^{T+K-1}} &= C_T + \mathbb{E}[z_{T+K}|\{I_t\}_{t=T}^{T+K-1}] \\
&= C_T + \mathbb{E}[z_{T+K}|I_{T+K-1}, I_{T+K-2}, q_{T+K-1}] \\
&= C_T + \mu_G\big[(1-\lambda)q_{T+K-1} + \lambda(1-q_{T+K-1})\big] + \mu_B\big[(1-\lambda)(1-q_{T+K-1}) + \lambda q_{T+K-1}\big] \\
&= C_T + \mu_G\big[q_{T+K-1} - \lambda(2q_{T+K-1}-1)\big] + \mu_B\big[1 - q_{T+K-1} + \lambda(2q_{T+K-1}-1)\big] \\
&= C_T + \big(\mu_G - \mu_B\big)\big[q_{T+K-1} - \lambda(2q_{T+K-1}-1)\big] + \mu_B
\end{aligned}$$

Note also that the fundamental price of the asset is $P_{T+K-1} = C_T + \mathbb{E}[z_{T+K}] = C_T + \big(\mu_G - \mu_B)p_G + \mu_B$. For any sequence $\{I_t\}_{t=T}^{T+K-1}$ this model generates over and underpricing as long as $q_{T+K-1} - \lambda(2q_{T+K-1}-1) \neq p_G$.

Recall that $q_t = \mathbb{P}(s_t = \text{Good}|I_t, I_{t-1}, q_{t-1})$. By Bayes rule, we have:

$$q_t = \frac{\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \text{Good}|I_{t-1}, q_{t-1})}{\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \text{Good}|I_{t-1}, q_{t-1}) + \mathbb{P}(I_t|s_t = \text{Bad}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \text{Bad}|I_{t-1}, q_{t-1})}$$

where:

$$\mathbb{P}(s_t = \text{Good}|I_{t-1}, q_{t-1}) = (1-\lambda)q_{t-1} + \lambda(1-q_{t-1})$$
$$= q_{t-1} - \lambda(2q_{t-1}-1)\eta_{t-1}\ .$$

The Bayesian update of q_t is:

$$q_t = \frac{\eta_{t-1}\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1})}{\eta_{t-1}\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1}) + (1-\eta_{t-1})\mathbb{P}(I_t|s_t = \text{Bad}, I_{t-1}, q_{t-1})}\ .$$

where $\eta_{t-1} = q_{t-1} - \lambda(2q_{t-1}-1)$. Given our model, we have to consider four particular states of the world, that we summarize in Table A.1.[14]

Table A.1: Belief updating rules

	$I_t = 1$	$I_t = -1$
$I_{t-1} = 1$	$\frac{\eta_{t-1}}{\eta_{t-1}(1-\gamma)+\gamma}$	$\frac{\eta_{t-1}(1-\pi_1)}{1-\pi_1(\gamma+\eta_{t-1}(1-\gamma))}$
$I_{t-1} = -1$	$\frac{\eta_{t-1}(1-\theta\pi_2)}{1-\pi_2(1+\eta_{t-1}(\theta-1))}$	$\frac{\eta_{t-1}\theta}{\eta_{t-1}(\theta-1)+1}$

In order to generate both under- and over-reaction of investors we need the four probabilities contained in Table A.1 to satisfy the following conditions $q_t^{1,1} \geq q_t^{-1,1} \geq q_t^{1,-1} \geq q_t^{-1,-1}$, which implies that investors attach a higher probability of being in a good state following two positive news. The condition above implies that:

$$\frac{\pi_2(\theta-1)}{\pi_2(\theta\gamma-1)+1-\gamma} \leq \pi_1 \leq \frac{\theta-1}{\theta\gamma-1} \quad \text{and} \quad \pi_2 \leq \frac{\gamma-1}{\theta\gamma-1}$$

Figures A.1, A.2 and A.3 contain the histograms of simulated prices for the correct pricing, underpricing and overpricing, respectively.

[14]The explicit resolution of the four cases can be found in Section E of the Online Appendix.

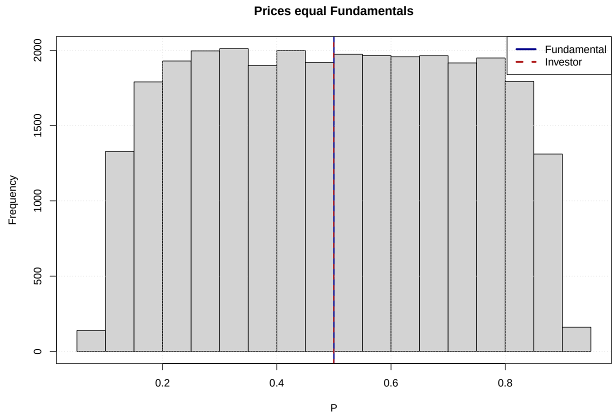


Figure A.1: Correct prices attain, on average, if $\pi_1 = \pi_2 = \lambda = \theta = \gamma = \frac{1}{2}$

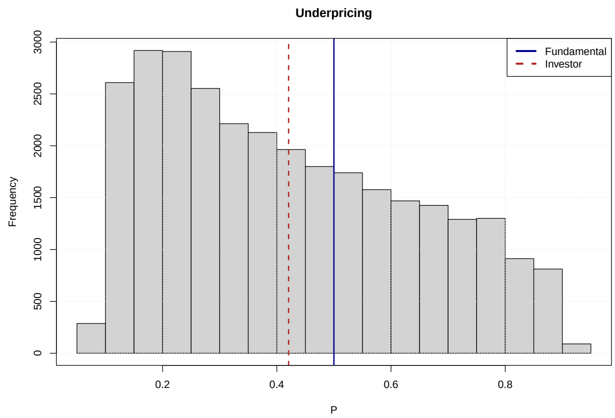


Figure A.2: Underpricing attains, on average, if $\pi_1 = \frac{2}{3}, \pi_2 = \lambda = \theta = \gamma = \frac{1}{2}$

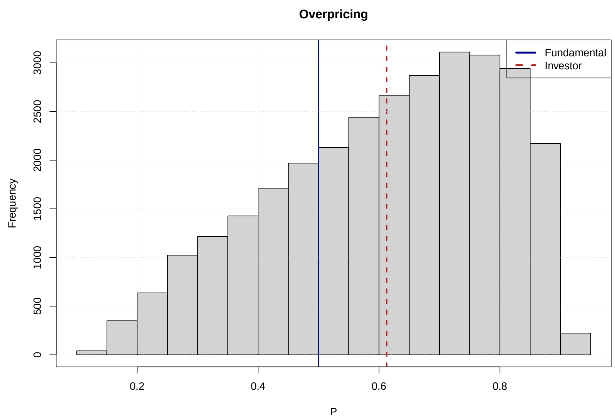


Figure A.3: Overpricing attains, on average, if $\pi_2 = \theta = \frac{2}{3}, \pi_1 = \lambda = \gamma = \frac{1}{2}$

Online Appendix

Directional Information and Momentum in Equity Returns

Luca Del Viva, Carlo Sala and André B.M. Souza

A Univariate Portfolio Sorting on Momentum

We compare the performance of the D-HML strategy with that of momentum. We follow Jegadeesh and Titman (1993), and Fama and French (1996) and form portfolios by sorting stocks into deciles based on their momentum (MOM), defined as the cumulative returns from month $t-12$ to $t-2$. Decile breakpoints are constructed based exclusively on stocks traded in the New York Stock Exchange (Barroso and Santa-Clara (2015)). For each month and decile, we build a value-weighted portfolio based on each stocks' lagged market capitalization and we indicate as winner-minus-loser (WML) the strategy that buys past winners and sells past losers.

As it can be noted from Panel A of Table A.2, the WML strategy produces average monthly returns of 1.03%. The average monthly return increases to 1.45% when controlling for the market, to 1.66% when controlling for the FF3 risk-factors, and to 1.44% with the FF5 specification. It becomes statistically insignificant when controlling for the Hou et al. (2020a) Q5 factors. The annualized Sharpe ratios of the WML strategy are 0.34 for the period from July 1932 to December 2022 and 0.43 for the period from July 1963 to December 2022, respectively. The WML strategy has a negative skewness of -2.12 and -0.80 for the July 1932 - December 2022 and July 1963 - December 2022 sample periods, respectively. The negative skewness and high kurtosis of the momentum strategy highlight the fact that it is prone to large extreme events. Moreover, the minimum and maximum returns of the WML strategy are of -77% and 52%, respectively.

Contrasting Panel A of Table A.2 with Panel A of Table 2, we note that the WML strategy has higher average abnormal returns than the D-HML strategy. However, this higher return is associated a substantially higher volatility. As a result, the Sharpe ratios of the WML are roughly half those of the D-HML strategy. Moreover, whereas the standard set of risk factors is able to explain up to 80% of the variation in momentum returns, they explain at most 25% of the variation in returns from the D-HML strategy.

About 14% of the stocks classified as momentum losers become winners, in contrast to the 10% of those in the bottom decile of PS, and about 11% of momentum winners become losers, in contrast to 9% of the high-PS stocks.

Momentum strategies generate high abnormal returns but experience severe, albeit rare, crashes along the way (Barroso and Santa-Clara (2015); Daniel and Moskowitz (2016)). The D-HML strategy, on the other hand, exhibits comparable risk adjusted returns with smaller volatility and substantially less frequent crashes.

B Market States

One of the primary reasons behind the occurrence of crashes in the momentum strategy is its poor performance during rebounds that follow bear markets (Daniel and Moskowitz (2016)). In this section we study how market states affect the performance of the D-HML and WML strategies. In particular, following Daniel and Moskowitz (2016) we study the abnormal returns of the two strategies during bear (BE) and bull (BU) markets. Bear markets are defined as periods where the past two-year cumulative return on the market is negative. To assess the impact of market states on our portfolio returns, we estimate the following regression:

$$r_{p,t} = \alpha_0 + \alpha_{\mathrm{BE}} \mathbb{1}_{\mathrm{BE}_{t-1}} + \tilde{r}_{m,t}(\beta_0 + \beta_{\mathrm{BE}} \mathbb{1}_{\mathrm{BE}_{t-1}}) + u_{p,t} \ , \tag{B.1}$$

where $r_{p,t}$ is the return on either the D-HML or WML portfolio in month t, $\mathbb{1}_{\mathrm{BE}_t}$ is an indicator variable representing past bear market states, and $\tilde{r}_{m,t}$ is the month t excess market return. The regression in Eq. (B.1) captures differences in excess returns and market exposures during bull and bear markets, respectively. α_0 and $\alpha_0 + \alpha_{\mathrm{BE}}$ capture the abnormal returns and β_0 and $\beta_0 + \beta_{\mathrm{BE}}$ are the market exposures. As in Daniel and Moskowitz (2016), we also study the performance of our strategies when the market rebounds following bear and bull markets. In particular, we estimate the following:

$$r_{p,t} = \alpha_0 + \alpha_{\text{BE}} \mathbb{1}_{\text{BE}_{t-1}} + \tilde{r}_{m,t} \big(\beta_0 + \beta_{\text{BE}} \mathbb{1}_{\text{BE}_{t-1}} + \beta^{+}_{\text{BE}} \mathbb{1}_{\text{BE}_{t-1}} \mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t}, \text{ and} \tag{B.2}$$

$$r_{p,t} = \alpha_0 + \alpha_{\text{BU}} \mathbb{1}_{\text{BU}_{t-1}} + \tilde{r}_{m,t} \big(\beta_0 + \beta_{\text{BU}} \mathbb{1}_{\text{BU}_{t-1}} + \beta^{+}_{\text{BU}} \mathbb{1}_{\text{BU}_{t-1}} \mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t} \ , \tag{B.3}$$

where $\mathbb{1}_{\text{BU}_{t-1}}$ is an indicator variable representing bull market states. The coefficients β^{+}_{BE} and β^{+}_{BU} capture the portfolio's performance during market rebounds following past bear or bull market periods. A negative β^{+}_{BE} or β^{+}_{BU} indicates a failure of the strategy in generating profits during market rebounds.

Panel A of Table A.3 shows that the WML strategy delivers positive and significant abnormal returns during bull markets (1.41% with t-stat 7.78). During bear markets, however, the WML strategy delivers negative and significant abnormal returns of -0.33% (1.41%-1.74%). The D-HML strategy, on the other hand, delivers positive and significant risk adjusted returns regardless of market state, with an additional premium of 0.69% during bear markets.

In line with Daniel and Moskowitz (2016), the WML strategy performs poorly during rebounds following bear markets (β^{+}_{BE}=-1.00 with t-stat -3.41). Although negative, the β^{+}_{BE} coefficients for the D-HML strategy is about a fifth of that of momentum and not statistically significant (-0.19 with t-stat=-1.36). In contrast to findings in Daniel and Moskowitz (2016), our WML strategy also fails when the market has positive returns following bull markets, with a negative and significant β^{+}_{BU} (-0.58 with t-stat=-2.3). The D-HML strategy has positive although not statistically significant premiums following bull markets (0.08% with t-stat=0.89).

Table A.4 in the Online Appendix carries out the same analysis for the sample period July 1963 - December 2022. Our previous findings are confirmed with sharper distinctions between the two strategies in the more recent sample period.

Overall, our results indicate that the D-HML strategy produces positive risk adjusted returns in both bear and bull markets and during market rebounds, and hence is not a directional bet on the market.

C Firm Characteristics

We analyse the following characteristic: i) A2ME is the ratio of total assets to market capitalization observed in December of the previous year. ii) AT is total assets. iii) ATO is net sales over lagged net operating assets. iv) BEME is the ratio between book value of equity to market value of equity observed in December of the previous year. v) BETA is the market beta estimated with daily data over a month. vi) C is the ratio of cash and short term investments to total assets. vii) CTO is the ratio of net sales to lagged total assets. viii) D2A is capital intensity calculated as the ratio of depreciation to total assets. ix) DPI2A is the ratio of changes in property, plant and equipment to lagged total assets. E2P is earnings to price ratio. x) FC2Y is the ratio between fixed costs to sales. xi) FREECF is the ratio of cash flows to book value of equity. xii) IMOM is the cumulative return from month $t-12$ to $t-7$. INVESTMENT is the yearly growth in total assets. xiii) IVOL is idiosyncratic volatility after the market model estimated over one month. xiv) LEV is financial leverage ratio. xv) LME is the previous month market capitalization. xvi) LTURNOVER is the last month volume to total shares outstanding. LTR is the cumulative return from month $t-36$ to $t-13$. xvii) MOM is the cumulative return from month $t-12$ to $t-2$. NOA is the ratio of net operating assets to lagged total assets. xviii) OA is the operating accruals calculated as in Sloan (1996). xix) OL is operating leverage calculated as the ratio costs of goods sold and selling, general and administrative expenses to total assets. xx) PCM is the price to cost margin calculated as the ratio of sales minus cost of goods sold to sales. xxi) PCTHIGH is the ratio of stock price at the end of the previous calendar month and the previous 52 week high price. PM is the ratio of operating income after depreciation to sales. xxii) PROF is gross profitability calculated as the ratio of gross profits to book value of equity. xxiii) Q is the Tobins'q. RNA is the return o net operating assets calculated as the ratio of operating income and net operating assets. xxiv) ROA is return on assets calculated as income before extraordinary items to lagged total assets. xxv) ROE is return on equity. xxvi) S2P is the ratio of sales to market capitalization observed in December of the previous year. xxvii) SGA2S is the ratio of selling, general and administrative expenses to sales. xxviii) SPREAD is the average daily difference between the lowest bid and highest ask scaled by the mid price. xxix) STR is the previous month return. xxx) SUV is the standard unexplained volume. Table A.5 contains the averages of characteristics for low and high PS stocks.

D Option-implied Skewness

To test whether the ranking generated by the directional signal is aligned with the traders in the option market, we construct two non-parametric option-implied indicators of skewness; one based upon the difference of out-of-the-money (OTM) call and put options (Bali et al. (2019)) and another one based upon the difference of OTM put options and (ATM) call options (Xing et al. (2010)). We base our main analysis on former, and use latter to test the robustness of the results. Among the many existing indicators of implied skewness, we choose the aforementioned approaches due to their empirically recognized robustness (Mixon (2011)).

Specifically, we infer the time series of individual stocks expected skewness from their implied volatilities which, in turn, are estimated by means of a basic kernel smoothing technique. The implied volatilities are obtained by first organizing the data by the log of days to expiration, and by "call-equivalent delta" (delta for a call, one plus delta for a put). Then, a smoothed volatility at each of the specified interpolation grid points is generated by means of a kernel smoother where, at each grid point on the surface, the smoothed implied volatilities are computed as weighted sums of option-implied volatilities. For the computations of our indicators we use American call (put) options market data with implied volatilties at delta level of 50 (-50) and 25 (-25) for the period 1996-2020. Finally, given the time window of the D-HML strategy, we only consider options with 30-day expiration.

Following Bali et al. (2019) the first indicator of implied skewness ($\mathrm{IS}^i_{t,\tau}$) is computed, for each day t and for each company i, as the difference between OTM call and OTM put implied volatilities:

$$\mathrm{IS}^i_{t,\tau} = \mathrm{IV}^i_{t,\tau}(0.25) - \mathrm{IV}^i_{t,\tau}(-0.25) \tag{D.4}$$

where $\mathrm{IV}^i_{t,\tau}(x)$ represents the time τ ahead implied-volatility (IV) with a delta equal to x.

As a robustness, we follow Xing et al. (2010) and propose another non-parametric option-implied indicator of skewness, this time defined as the difference between OTM put and ATM call implied volatilities:

$$\mathrm{IS}^i_{t,\tau} = \mathrm{IV}^i_{t,\tau}(-0.25) - \mathrm{IV}^i_{t,\tau}(0.50) \tag{D.5}$$

where, again, $\mathrm{IV}^i_{t,\tau}(x)$ represents the time τ ahead implied-volatility (IV) with a delta equal to x.

E A Model of Investors Biased Expectations

Recall that $q_t = \mathbb{P}(s_t = \mathrm{Good}|I_t, I_{t-1}, q_{t-1})$. By Bayes rule, we have:

$$q_t = \frac{\mathbb{P}(I_t|s_t = \mathrm{Good}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \mathrm{Good}|I_{t-1}, q_{t-1})}{\mathbb{P}(I_t|s_t = \mathrm{Good}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \mathrm{Good}|I_{t-1}, q_{t-1}) + \mathbb{P}(I_t|s_t = \mathrm{Bad}, I_{t-1}, q_{t-1})\mathbb{P}(s_t = \mathrm{Bad}|I_{t-1}, q_{t-1})}$$

where:

$$\mathbb{P}(s_t = \text{Good}|I_{t-1}, q_{t-1}) = (1-\lambda)q_{t-1} + \lambda(1-q_{t-1})$$
$$= q_{t-1} - \lambda(2q_{t-1}-1)\eta_{t-1} \ .$$

The Bayesian update of q_t is:

$$q_t = \frac{\eta_{t-1}\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1})}{\eta_{t-1}\mathbb{P}(I_t|s_t = \text{Good}, I_{t-1}, q_{t-1}) + (1-\eta_{t-1})\mathbb{P}(I_t|s_t = \text{Bad}, I_{t-1}, q_{t-1})} \ .$$

where $\eta_{t-1} = q_{t-1} - \lambda(2q_{t-1} - 1)$.

We have to consider four particular states of the world. If the shocks at time t and $t-1$ are both positive (i.e., $I_t = I_{t-1} = 1$) then the investor updates q_t as follows:

$$q_t^{1,1} = \frac{\eta_{t-1}\pi_1}{\eta_{t-1}\pi_1 + (1-\eta_{t-1})\gamma\pi_1} = \frac{\eta_{t-1}}{\eta_{t-1}(1-\gamma)+\gamma}$$

If the investor receives two consecutive negative news so that shocks at time t and $t-1$ are both negative (i.e., $I_t = I_{t-1} = -1$) then the investor updates q_t as follows:

$$q_t^{-1,-1} = \frac{\eta_{t-1}\theta\pi_2}{\eta_{t-1}\theta\pi_2 + (1-\eta_{t-1})\pi_2} = \frac{\eta_{t-1}\theta}{\eta_{t-1}(\theta-1)+1}$$

If a negative news received at time $t-1$ is followed by a positive news at time t (i.e., $I_{t-1} = -1$ and $I_t = 1$) then the investor updates q_t as follows:

$$q_t^{-1,1} = \frac{\eta_{t-1}(1-\theta\pi_2)}{\eta_{t-1}(1-\theta\pi_2) + (1-\eta_{t-1})(1-\pi_2)} = \frac{\eta_{t-1}(1-\theta\pi_2)}{1-\pi_2(1+\eta_{t-1}(\theta-1))}$$

Finally, if a positive news received at time $t-1$ is followed by a negative news at time t (i.e., $I_{t-1} = 1$ and $I_t = -1$) then the investor updates q_t as:

$$q_t^{1,-1} = \frac{\eta_{t-1}(1-\pi_1)}{\eta_{t-1}(1-\pi_1) + (1-\eta_{t-1})(1-\gamma\pi_1)} = \frac{\eta_{t-1}(1-\pi_1)}{1-\pi_1(\gamma+\eta_{t-1}(1-\gamma))}$$

Table A.1:

Estimation of the Probability Score

The table reports the coefficients obtained in the estimation of our probability score using the full sample. In particular, using the full sample of data we run the following panel regression $r^+_{i,t} = \delta_0 + \delta_1 \, IV_{i,t-1} + \delta_2 \, r_{m,t-1} + \beta^m_+ P^m_{i,t-1} + \beta^w_+ P^w_{i,t-1} + \beta^d_+ P^d_{i,t-1} + \beta^m_- N^m_{i,t-1} + \beta^w_- N^w_{i,t-1} + \beta^d_- N^d_{i,t-1} + u_{i,t}$. The dependent variable is constructed as $r^+_{i,t} = \mathbb{1}(r_{i,t} > 0)$, taking the value of one if the return of stock i in month t is strictly positive and zero otherwise. $IV_{i,t-1}$ is the lagged idiosyncratic variance, constructed as the variance of the residuals after estimating the market model with daily returns over a month, $r_{m,t-1}$ is the lagged market return, and $\{P^j_{i,t-1}, N^j_{i,t-1}\}$ for $j \in \{d, w, m\}$ represent the duration of the current run of daily, weekly, and monthly positive and negative returns, respectively. Robust t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	(1)	(2)	(3)	(4)
Constant	0.514***	0.509***	0.507***	0.506***
	(1110.770)	(678.581)	(568.328)	(495.562)
$IV_{i,t-1}$	−0.420***	−0.423***	−0.423***	−0.424***
	(−2.662)	(−2.670)	(−2.671)	(−2.684)
$r_{m,t-1}$	0.436***	0.453***	0.457***	0.458***
	(64.728)	(58.314)	(57.592)	(57.349)
$P^m_{i,t-1}$		0.002***	0.002***	0.003***
		(7.037)	(7.345)	(11.348)
$P^w_{i,t-1}$			0.001**	0.002***
			(2.301)	(5.782)
$P^d_{i,t-1}$				−0.010***
				(−30.050)
$N^m_{i,t-1}$		0.004***	0.004***	0.002***
		(12.494)	(11.727)	(7.940)
$N^w_{i,t-1}$			0.001***	0.001***
			(5.243)	(2.873)
$N^d_{i,t-1}$				0.011***
				(34.377)
R^2 (x100)	0.195	0.203	0.204	0.368

Table A.2:

Value-Weighted Portfolios sorted on MOM

Panel A reports raw and abnormal returns (alphas) of the winner minus loser portfolio obtained as the difference between the value-weighted winners and losers (WML) portfolios constructed on momentum (MOM). Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); (v) the Hou et al. (2020a) 5-factor model (Q5); v) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurstosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtotis of the portfolio returns, respectively. Panel B reports the FF3 abnormal returns for each value weighted portfolio built on deciles of MOM. Max and Min indicate the maximum and minimum returns observed in each portfolio, respectively. Avg. Pos. Ret. is the average over time of the percentage of positive returns in each decile. Avg. Winners and Avg. Losers are the average over time of the percentage of stocks in each portfolio that are also in the top and bottom decile of future returns, respectively. Avg. Market Cap is the average over time of the median market capitalization (in thousands) of firms in each portfolio. The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

Panel A. Raw and abnormal returns (alphas) of the winners minus losers portfolios (WML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	1.03***	1.45***	1.66***	1.44***	0.41***	0.17	0.31
	(4.57)	(8.08)	(8.75)	(4.52)	(3.47)	(0.50)	(1.58)
MKT		−0.54***	−0.38***	−0.34***	−0.07*	−0.19*	−0.07
		(−4.92)	(−4.72)	(−3.33)	(−1.96)	(−1.92)	(−1.58)
SMB			−0.21	−0.02	−0.05		
			(−1.59)	(−0.09)	(−0.78)		
HML			−0.69***	−0.87***	−0.06		
			(−5.15)	(−3.80)	(−0.79)		
RMW				0.36	0.08		
				(1.16)	(1.05)		
CMA				0.63*	0.10		
				(1.72)	(0.76)		
UMD					1.50***		1.47***
					(46.59)		(35.08)
STR					−0.15**		−0.16**
					(−2.51)		(−2.51)
LTR					−0.05		−0.03
					(−0.62)		(−0.35)
R_{ME}						0.35*	−0.01
						(1.68)	(−0.18)
R_{IA}						−0.20	−0.04
						(−0.81)	(−0.35)
R_{ROE}						1.15***	0.06
						(4.69)	(0.86)
R_{EG}						0.73***	0.17
						(3.00)	(1.23)
T	1,086	1,086	1,086	714	714	672	672
R^2	0.02	0.14	0.23	0.14	0.80	0.29	0.80
Vol	7.85	7.85	7.85	7.38	7.38	7.53	7.53
SR	0.34	0.34	0.34	0.43	0.43	0.41	0.41
Skew	-2.12	-2.12	-2.12	-0.80	-0.80	-0.80	-0.80
Kurtosis	21.24	21.24	21.24	11.69	11.69	11.42	11.42

Panel B. Abnormal returns for each decile portfolios built on MOM

	1	2	3	4	5	6	7	8	9	10	10 − 1
Alpha	−1.13***	−0.47***	−0.28***	−0.09	−0.10*	−0.00	0.11**	0.21***	0.24***	0.53***	1.66***
	(-8.18)	(-5.08)	(-3.48)	(-1.33)	(-1.78)	(-0.09)	(2.21)	(4.03)	(3.71)	(6.01)	(8.75)
MKT	1.41***	1.23***	1.11***	1.06***	1.01***	1.01***	0.97***	0.95***	0.98***	1.03***	−0.38***
	(24.85)	(27.69)	(32.27)	(32.88)	(35.81)	(42.20)	(48.56)	(38.79)	(35.65)	(29.18)	(-4.72)
SMB	0.50***	0.14**	−0.02	−0.06*	−0.06	−0.05	−0.10***	−0.07**	−0.02	0.29***	−0.21
	(5.42)	(2.05)	(-0.19)	(-1.69)	(-1.56)	(-1.32)	(-2.74)	(-2.26)	(-0.45)	(5.21)	(-1.59)
HML	0.39***	0.38***	0.32***	0.24***	0.20***	0.16***	0.06*	0.01	−0.05	−0.29***	−0.69***
	(4.23)	(5.74)	(5.90)	(7.25)	(4.06)	(4.98)	(1.92)	(0.17)	(-1.15)	(-5.30)	(-5.15)
Vol	9.62	7.63	6.6	5.99	5.6	5.5	5.17	5.1	5.37	6.3	7.85
SR	0.09	0.31	0.38	0.48	0.48	0.53	0.59	0.64	0.63	0.71	0.34
Skew	1.98	2.53	2.21	2.02	1.91	1.38	0.52	0.27	-0.15	-0.41	-2.12
Max	0.94	0.8	0.67	0.62	0.61	0.49	0.35	0.33	0.36	0.3	0.52
Min	-0.56	-0.4	-0.34	-0.33	-0.32	-0.31	-0.26	-0.25	-0.27	-0.28	-0.77
Avg. Pos. Ret.	0.43	0.47	0.49	0.5	0.51	0.52	0.53	0.53	0.53	0.52	–
Avg. Winners	0.14	0.1	0.08	0.08	0.07	0.07	0.07	0.08	0.09	0.13	–
Avg. Losers	0.19	0.11	0.08	0.07	0.06	0.06	0.06	0.06	0.07	0.11	–
Avg. Market Cap	36,053	109,215	166,811	216,172	252,163	284,614	297,049	315,814	317,509	196,754	–

Table A.3:

Directional signal, Momentum and Market States

Each month value-weighted decile portfolios are constructed according to deciles based on the directional signal (PS) and the original momentum (MOM). D-HML is the directional high minus low portfolio. WML is the momentum winners minus losers strategy. For each strategy we estimate $r_{p,t} = \alpha_0 + \alpha_{\mathrm{BE}} \mathbb{1}_{\mathrm{BE}_{t-1}} + \tilde{r}_{m,t}\big(\beta_0 + \beta_{\mathrm{BE}} \mathbb{1}_{\mathrm{BE}_{t-1}} + \beta^{+}_{\mathrm{BE}} \mathbb{1}_{\mathrm{BE}_{t-1}} \mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t}$ and $r_{p,t} = \alpha_0 + \alpha_{\mathrm{BU}} \mathbb{1}_{\mathrm{BU}_{t-1}} + \tilde{r}_{m,t}\big(\beta_0 + \beta_{\mathrm{BU}} \mathbb{1}_{\mathrm{BU}_{t-1}} + \beta^{+}_{\mathrm{BU}} \mathbb{1}_{\mathrm{BU}_{t-1}} \mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t}$ where $r_{p,t}$ is the return of a portfolio in month t; $\mathbb{1}_{\mathrm{BE}_{t-1}}$ and $\mathbb{1}_{\mathrm{BU}_{t-1}}$ are binary dummy variables taking the value 1 if the past two-year cumulative market return up to month $t-1$ is negative (bear market) and positive (bull market), respectively; $\tilde{r}_{m,t}$ is the excess market return; $\mathbb{1}\{\tilde{r}_{m,t} > 0\}$ is a dummy variable taking the value 1 if the contemporaneous (not ex ante) excess market return is positive. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively. The sample period is from July 1932 to December 2022.

Panel A. Portfolio returns and ex-ante bear market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.04 (−0.70)	0.62*** (8.47)	0.66*** (6.58)	−0.70*** (−5.26)	0.71*** (6.41)	1.41*** (7.78)
α_{BE}	−0.49 (−1.54)	0.20 (0.86)	0.69* (1.67)	1.18** (2.26)	−0.56** (−2.14)	−1.74*** (−2.58)
β_0	0.96*** (51.58)	1.03*** (57.44)	0.08*** (2.89)	1.34*** (22.45)	1.22*** (33.90)	−0.12 (−1.43)
β_{BE}	0.05 (1.17)	−0.02 (−0.72)	−0.07 (−1.30)	0.56*** (4.78)	−0.43*** (−6.34)	−0.99*** (−5.84)
R^2	0.84	0.88	0.01	0.74	0.74	0.24

Panel B. Portfolio returns, ex-ante bear and contemporaneous positive market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.04 (−0.68)	0.62*** (8.27)	0.66*** (6.76)	−0.70*** (−5.27)	0.71*** (6.39)	1.41*** (7.65)
α_{BE}	−0.90** (−1.98)	0.34 (1.14)	1.24** (2.13)	−0.62 (−0.79)	0.56 (1.56)	1.18 (1.15)
β_0	0.96*** (52.10)	1.03*** (57.47)	0.08*** (2.88)	1.34*** (22.46)	1.22*** (33.52)	−0.12 (−1.45)
β_{BE}	−0.05 (−0.61)	0.01 (0.11)	0.05 (0.51)	0.17 (1.14)	−0.18* (−1.93)	−0.35* (−1.65)
β^{+}_{BE}	0.14 (1.46)	−0.05 (−0.64)	−0.19 (−1.36)	0.62*** (3.21)	−0.38*** (−3.47)	−1.00*** (−3.41)
R^2	0.84	0.88	0.01	0.74	0.74	0.25

Panel C. Portfolio returns, ex-ante bull and contemporaneous positive market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.53* (−1.75)	0.82*** (3.50)	1.35*** (3.45)	0.48 (0.93)	0.15 (0.65)	−0.33 (−0.51)
α_{BU}	0.44 (1.38)	−0.38 (−1.47)	−0.82* (−1.94)	−1.73*** (−3.00)	0.97*** (3.36)	2.71*** (3.59)
β_0	1.00*** (25.19)	1.01*** (36.79)	0.01 (0.17)	1.90*** (18.62)	0.79*** (13.34)	−1.11*** (−7.75)
β_{BU}	−0.06 (−1.28)	−0.03 (−0.71)	0.03 (0.47)	−0.73*** (−5.67)	0.56*** (6.54)	1.29*** (7.00)
β^{+}_{BU}	0.03 (0.56)	0.11* (1.95)	0.08 (0.89)	0.33* (1.91)	−0.25** (−2.55)	−0.58** (−2.43)
R^2	0.84	0.88	0.01	0.74	0.74	0.24

Table A.4:

Directional signal, Momentum and Market States

Each month value-weighted decile portfolios are constructed according to deciles based on the directional signal (PS) and the original momentum (MOM). D-HML is the directional high minus low portfolio. WML is the momentum winners minus losers strategy. For each strategy we estimate $r_{p,t} = \alpha_0 + \alpha_{\mathrm{BE}}\mathbb{1}_{\mathrm{BE}_{t-1}} + \tilde{r}_{m,t}\big(\beta_0 + \beta_{\mathrm{BE}}\mathbb{1}_{\mathrm{BE}_{t-1}} + \beta^{+}_{\mathrm{BE}}\mathbb{1}_{\mathrm{BE}_{t-1}}\mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t}$ and $r_{p,t} = \alpha_0 + \alpha_{\mathrm{BU}}\mathbb{1}_{\mathrm{BU}_{t-1}} + \tilde{r}_{m,t}\big(\beta_0 + \beta_{\mathrm{BU}}\mathbb{1}_{\mathrm{BU}_{t-1}} + \beta^{+}_{\mathrm{BU}}\mathbb{1}_{\mathrm{BU}_{t-1}}\mathbb{1}\{\tilde{r}_{m,t} > 0\}\big) + u_{p,t}$ where $r_{p,t}$ is the return of a portfolio in month t; $\mathbb{1}_{\mathrm{BE}_{t-1}}$ and $\mathbb{1}_{\mathrm{BU}_{t-1}}$ are binary dummy variables taking the value 1 if the past two-year cumulative market return up to month $t-1$ is negative (bear market) and positive (bull market), respectively; $\tilde{r}_{m,t}$ is the excess market return; $\mathbb{1}\{\tilde{r}_{m,t} > 0\}$ is a dummy variable taking the value 1 if the contemporaneous (not ex ante) excess market return is positive. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively. The sample period is July 1963 to December 2022.

Panel A. Portfolio returns and ex-ante bear market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.01 (−0.08)	0.82*** (10.78)	0.83*** (5.96)	−0.68*** (−3.82)	0.81*** (5.46)	1.49*** (6.12)
α_{BE}	−0.46 (−1.39)	0.41 (1.34)	0.87 (1.58)	1.83* (1.95)	−0.51 (−1.42)	−2.34* (−1.92)
β_0	0.95*** (38.55)	1.03*** (44.92)	0.08** (2.33)	1.35*** (19.27)	1.23*** (27.93)	−0.11 (−1.18)
β_{BE}	−0.01 (−0.06)	0.07 (0.96)	0.07 (0.54)	0.56** (2.33)	−0.37*** (−4.34)	−0.93*** (−3.01)
R^2	0.82	0.85	0.02	0.65	0.72	0.12

Panel B. Portfolio returns, ex-ante bear and contemporaneous positive market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.01 (−0.08)	0.82*** (10.47)	0.83*** (6.34)	−0.68*** (−3.80)	0.81*** (5.52)	1.49*** (6.14)
α_{BE}	0.58 (0.95)	0.45 (0.97)	−0.12 (−0.14)	0.38 (0.31)	−0.06 (−0.12)	−0.45 (−0.27)
β_0	0.95*** (38.02)	1.03*** (45.31)	0.08** (2.27)	1.35*** (19.14)	1.23*** (28.09)	−0.11 (−1.16)
β_{BE}	0.17 (1.39)	0.08 (0.71)	−0.1 (−0.54)	0.31** (2.28)	−0.29*** (−2.83)	−0.6*** (−3.01)
β^{+}_{BE}	−0.39* (−1.67)	−0.02 (−0.09)	0.38 (0.99)	0.55 (0.90)	−0.17 (−0.71)	−0.72 (−0.85)
R^2	0.82	0.85	0.03	0.65	0.72	0.12

Panel C. Portfolio returns, ex-ante bull and contemporaneous positive market

	PS			MOM		
	1	10	10 − 1	1	10	10 − 1
α_0	−0.47 (−1.46)	1.23*** (3.89)	1.69*** (3.31)	1.15 (1.25)	0.3 (0.97)	−0.85 (−0.67)
α_{BU}	0.51 (1.50)	−0.63* (−1.84)	−1.14** (−2.08)	−2.29** (−2.34)	0.76** (2.02)	3.05** (2.25)
β_0	0.94*** (10.69)	1.1*** (16.54)	0.16 (1.31)	1.91*** (8.18)	0.86*** (12.10)	−1.04*** (−3.41)
β_{BU}	0.02 (0.21)	−0.14* (−1.76)	−0.16 (−1.20)	−0.71*** (−2.75)	0.45*** (4.11)	1.16*** (3.42)
β^{+}_{BU}	−0.03 (−0.41)	0.14* (1.90)	0.17 (1.48)	0.28 (1.51)	−0.16 (−1.43)	−0.44* (−1.68)
R^2	0.82	0.85	0.02	0.65	0.72	0.12

Table A.5:

Directional signal and Firm Characteristics

This Table contains the average firm characteristics for the bottom and top deciles built on PS. i) A2ME is the ratio of total assets to market capitalization observed in December of the previous year. ii) AT is total assets. iii) ATO is net sales over lagged net operating assets. iv) BEME is the ratio between book value of equity to market value of equity observed in December of the previous year. v) BETA is the market beta estimated with daily data over a month. vi) C is the ratio of cash and short term investments to total assets. vii) CTO is the ratio of net sales to lagged total assets. viii) D2A is capital intensity calculated as the ratio of depreciation to total assets. ix) DPI2A is the ratio of changes in property, plant and equipment to lagged total assets. x) E2P is earnings to price ratio. xi) FC2Y is the ratio between fixed costs to sales. xii) FREECF is the ratio of cash flows to book value of equity. xiii) IMOM is the cumulative return from month $t-12$ to $t-7$. xiv) INVESTMENT is the yearly growth in total assets. xv) IVOL is idiosyncratic volatility after the market model estimated over one month. xvi) LEV is financial leverage ratio. xvii) LME is the previous month market capitalization. xviii) LTURNOVER is the last month volume to total shares outstanding. xix) LTR is the cumulative return from month $t-36$ to $t-13$. xx) MOM is the cumulative return from month $t-12$ to $t-2$. xxi) NOA is the ratio of net operating assets to lagged total assets. xxii) OA is the operating accruals calculated as in Sloan (1996). xxiii) OL is operating leverage calculated as the ratio costs of goods sold and selling, general and administrative expenses to total assets. xxiv) PCM is the price to cost margin calculated as the ratio of sales minus cost of goods sold to sales. xxv) PCTHIGH is the ratio of stock price at the end of the previous calendar month and the previous 52 week high price. xxvi) PM is the ratio of operating income after depreciation to sales. xxvii) PROF is gross profitability calculated as the ratio of gross profits to book value of equity. xxviii) Q is the Tobins'q. RNA is the return o net operating assets calculated as the ratio of operating income and net operating assets. xxix) ROA is return on assets calculated as income before extraordinary items to lagged total assets. xxx) ROE is return on equity. xxxi) S2P is the ratio of sales to market capitalization observed in December of the previous year. xxxii) SGA2S is the ratio of selling, general and administrative expenses to sales. xxxiii) SPREAD is the average daily difference between the lowest bid and highest ask scaled by the mid price. xxxiv) STR is the previous month return. SUV is the standard unexplained volume.

	Low PS	High PS	Diff.		Low PS	High PS	Diff.
A2ME	3.65*** (97.65)	3.12	−0.53*** (−10.01)	AT	5.42*** (1060.28)	5.59	0.17*** (23.28)
ATO	2.40*** (239.28)	2.46	0.05*** (3.89)	BEME	0.83*** (157.39)	0.78	−0.04*** (−5.62)
BETA	0.85*** (225.87)	0.76	−0.09*** (−17.20)	C	0.15*** (349.20)	0.15	−0.00 (−0.30)
CTO	1.32*** (38.50)	1.40	0.08* (1.75)	D2A	0.04*** (430.25)	0.04	−0.00*** (−8.28)
DPI2A	0.16*** (3.45)	0.20	0.05 (0.70)	E2P	−0.05*** (−16.21)	0.02	0.07*** (14.85)
FC2Y	0.68*** (10.78)	0.73	0.06 (0.67)	FREECF	−0.20** (−2.19)	−0.09	0.11 (0.86)
IMOM	0.09*** (82.21)	0.09	0.01*** (4.42)	INVESTMENT	0.39*** (7.51)	0.48	0.09 (1.27)
IVOL	0.04*** (500.10)	0.03	−0.01*** (−118.16)	LEV	0.34*** (131.53)	0.34	−0.00 (−1.19)
LME	11.92*** (2432.21)	12.15	0.23*** (33.53)	LTR	0.37*** (125.19)	0.44	0.07*** (15.63)
LTURNOVER	1.07*** (150.98)	1.11	0.04*** (3.70)	MOM	0.16*** (104.35)	0.15	−0.01*** (−5.94)
NOA	0.68*** (15.77)	0.74	0.06 (0.99)	OA	−0.02*** (−5.16)	−0.02	0.00 (0.13)
OL	1.11*** (488.24)	1.09	−0.02*** (−5.33)	PCM	−1.35*** (−6.43)	−0.71	0.64** (2.17)
PCTHIGH	0.74*** (1471.64)	0.73	−0.00*** (−3.31)	PM	−1.93*** (−8.96)	−1.31	0.63** (2.07)
PROF	0.82*** (9.25)	0.66	−0.16 (−1.30)	Q	2.25*** (62.72)	2.35	0.10** (2.06)
RNA	0.12*** (35.55)	0.15	0.03*** (5.84)	ROA	−0.01** (−2.47)	0.01	0.02*** (5.06)
ROE	−0.03 (−0.61)	0.04	0.07 (0.95)	S2P	2.65*** (182.98)	2.22	−0.43*** (−21.19)
SGA2S	0.54*** (13.00)	0.58	0.03 (0.57)	SPREAD	0.05*** (406.85)	0.04	−0.01*** (−58.25)
STR	0.01*** (14.06)	−0.01	−0.02*** (−29.99)	SUV	−1.32*** (−644.46)	−1.34	−0.02*** (−7.54)

Table A.6:

List of Anomalies

The table contains the list of asset pricing anomalies used in the construction of the *Net* measure as in Engelberg et al. (2018). Rank in each anomaly is obtained from Chen and Zimmermann (2022).

List of Anomalies (Chen and Zimmermann (2022))							
1	AbnormalAccruals	46	DelEqu	91	IntMom	136	ProbInformedTrading
2	Accruals	47	DelFINL	92	Investment	137	PS
3	Activism1	48	DelLTI	93	InvestPPEInv	138	RD
4	Activism2	49	DelNetFin	94	InvGrowth	139	RDAbility
5	AdExp	50	dNoa	95	IO_ShortInterest	140	RDcap
6	AgeIPO	51	DolVol	96	Leverage	141	RDS
7	AM	52	EarningsConsistency	97	LRreversal	142	realestate
8	AnalystRevision	53	EarningsForecastDisparity	98	MaxRet	143	ResidualMomentum
9	AnalystValue	54	EarningsStreak	99	MeanRankRevGrowth	144	retConglomerate
10	AnnouncementReturn	55	EarningsSurprise	100	Mom6m	145	ReturnSkew
11	AOP	56	EarnSupBig	101	Mom6mJunk	146	ReturnSkew3F
12	AssetGrowth	57	EBM	102	Mom12m	147	REV6
13	Beta	58	EntMult	103	Mom12mOffSeason	148	RevenueSurprise
14	BetaFP	59	EP	104	MomOffSeason	149	roaq
15	BetaLiquidityPS	60	EquityDuration	105	MomOffSeason06YrPlus	150	RoE
16	BetaTailRisk	61	ExclExp	106	MomOffSeason11YrPlus	151	sfe
17	betaVIX	62	FEPS	107	MomOffSeason16YrPlus	152	ShareIss1Y
18	BidAskSpread	63	fgr5yr	108	MomSeason	153	ShareIss5Y
19	BM	64	FirmAge	109	MomSeason06YrPlus	154	ShortInterest
20	Bmdec	65	FirmAgeMom	110	MomSeason11YrPlus	155	size
21	BookLeverage	66	ForecastDispersion	111	MomSeason16YrPlus	156	skew1
22	BPEBM	67	FR	112	MomSeasonShort	157	SmileSlope
23	BrandInvest	68	Frontier	113	Mrreversal	158	SP
24	Cash	69	GP	114	NetDebtFinance	159	std_turn
25	CashProd	70	GrAdExp	115	NetDebtPrice	160	STreversal
26	CBOperProf	71	grcapx	116	NetEquityFinance	161	tang
27	CF	72	grcapx3y	117	NetPayoutYield	162	Tax
28	cfp	73	GrLTNOA	118	NOA	163	TotalAccruals
29	ChangeInRecommendation	74	GrSaleToGrInv	119	NumEarnIncrease	164	VarCF
30	ChAssetTurnover	75	GrSaleToGrOverhead	120	OperProf	165	VolMkt
31	ChEQ	76	Herf	121	OperProfRD	166	VolSD
32	ChInv	77	HerfAsset	122	OPLeverage	167	VolumeTrend
33	ChInvIA	78	HerfBE	123	OptionVolume1	168	XFIN
34	ChNNCOA	79	High52	124	OptionVolume2	169	zerotrade
35	ChNWC	80	hire	125	OrderBacklog	170	zerotradeAlt1
36	ChTax	81	IdioRisk	126	OrderBacklogChg	171	zerotradeAlt12
37	CompEquIss	82	IdioVol3F	127	OrgCap		
38	CompositeDebtIssuance	83	IdioVolAHT	128	PayoutYield		
39	CoskewACX	84	Illiquidity	129	PctAcc		
40	Coskewness	85	IndMom	130	PctTotAcc		
41	CustomerMomentum	86	IndRetBig	131	PredictedFE		
42	DelBreadth	87	IntanBM	132	Price		
43	DelCOA	88	IntanCFP	133	PriceDelayRsq		
44	DelCOL	89	IntanEP	134	PriceDelaySlope		
45	DelDRC	90	IntanSP	135	PriceDelayTstat		

Table A.7:

Probability Score and option-implied Skewness

The table reports abnormal returns for the Low (1st decile), High (10th decile) and directional high minus low (D-HML) portfolios constructed on the probability score PS. Abnormal returns (alphas) are generated using i) the capital asset pricing model of Sharpe (1964), Lintner (1965) and Mossin (1966), ii) the Fama and French (1993) three factor model (FF3) iii) and the Fama and French (2015) five factor model (FF5). Each asset pricing model is augmented with the options implied skewness factor (ISK) calculated as the difference in the options implied skewness of overly pessimistic minus overly optimistic stocks. T is the number of months in the sample period. Vol, SR, Skew and Kurtosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtosis of the portfolio returns. The sample period extends from February 1996 to December 2020. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	CAPM	CAPM + ISK	FF3	FF3 + ISK	FF5	FF5 + ISK
Alpha	0.95^{***}	0.17	0.96^{***}	0.07	0.98^{***}	0.08
	(3.96)	(0.50)	(3.98)	(0.20)	(3.83)	(0.21)
MKT	0.07	0.09	0.06	0.09	0.06	0.09
	(0.91)	(1.33)	(0.94)	(1.50)	(0.95)	(1.42)
SMB			−0.05	−0.09	−0.08	-0.14^{*}
			(−0.65)	(−1.10)	(−0.98)	(−1.76)
HML			-0.23^{*}	-0.26^{**}	−0.22	−0.22
			(−1.93)	(−2.31)	(−1.37)	(−1.40)
RMW					−0.06	−0.11
					(−0.49)	(−0.98)
CMA					0.05	0.04
					(0.29)	(0.24)
ISK		-0.37^{***}		-0.42^{***}		-0.45^{***}
		(−2.91)		(−3.03)		(−3.10)
T	299	299	299	299	299	299
R^2	0.06	0.10	0.09	0.13	0.09	0.13
Vol	3.79	3.79	3.79	3.79	3.79	3.79
SR	0.75	0.75	0.75	0.75	0.75	0.75
Skew	0.95	0.95	0.95	0.95	0.95	0.95
Kurtosis	5.70	5.70	5.70	5.70	5.70	5.70

Table A.8:

Portfolio Analysis With Competing Asset Pricing Models

The table reports abnormal returns (alphas) of the directional high minus low signal portfolios (D-HML) obtained as difference between high-PS and low-PS. Abnormal returns are estimated with the following models: 1) the Fama and French (1993) 3-factor models (FF3); 2) the FF3 model augmented by the momentum, short-term reversal and long-term reversal factors (FF3'); 3) the FF3' model augmented by the Pástor and Stambaugh (2003) liquidity factor (FF3"); 4) the Fama and French (2015) 5-factor model augmented by the momentum, short-term reversal, long-term reversal and liquidity factor (FF5'); 5) the Hou et al. (2015) 4-factor model (Q4); 6) the Hou et al. (2020a) 5-factor model augmented by the momentum, short-term reversal, long-term reversal and liquidity factor (Q5'); 7) the Stambaugh and Yuan (2017) model (SY (2017)) with the MGMT and PERF mispricing factors; 8) the Daniel et al. (2020) model (DHS (2020)) with the FIN and PEAD behavioral factors; 9) the Asness et al. (2019) model (AFP (2019)) with the quality minus junk QMJ factor. Models (10) to (15) are obtained augmenting the FF5 model with: i) the Bali et al. (2017) FMAX factor; ii) the Frazzini and Pedersen (2014) betting-against-beta (BAB) factor; iii) the Asness et al. (2000) betting-against-correlation (BAC) factor; iv) the betting-against-volatility (BAV) factor; v) the Atilgan et al. (2020) left-tail momentum (LTM) factor; and vi) the idiosyncratic volatility IVOL factor estimated from the variance of residuals after the Fama and French (1993) 3-factor model estimated on daily returns over the previous 2 months. Newey-West adjusted t-statistics are given in parentheses. The sample period is from July 1932 to December 2022 but varies depending on the asset pricing specification. T is the total number of months in the sample period and R^2 is the coefficient of determination. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	(1) FF3	(2) FF3'	(3) FF3"	(4) FF5'	(5) Q4	(6) Q5'	(7) SY (2017)	(8) DHS (2020)	(9) AFP (2019)	(10) FMAX	(11) BAB	(12) BAC	(13) BAV	(14) LTM	(15) IVOL
Alpha	0.82***	0.70***	0.93***	0.96***	1.11***	0.93***	1.16***	1.14***	0.93***	0.91***	1.06***	1.04***	1.02***	1.01***	0.99***
	(6.44)	(5.04)	(5.44)	(5.25)	(6.48)	(4.90)	(6.29)	(6.18)	(6.34)	(6.15)	(7.30)	(7.26)	(7.33)	(6.22)	(6.44)
MKT	0.03	−0.03	−0.04	−0.04	0.05	−0.03	0.04	0.04	0.10**	0.16***	0.06	0.05	0.05	0.05	0.10**
	(0.95)	(−1.02)	(−0.90)	(−1.09)	(1.07)	(−0.82)	(0.71)	(0.82)	(2.35)	(3.83)	(1.36)	(1.24)	(1.30)	(1.09)	(2.52)
SMB	0.05	−0.01	−0.03	−0.06			0.04	0.04	0.07	0.16**	0.05	0.05	0.00	0.04	0.11
	(1.04)	(−0.19)	(−0.57)	(−1.06)			(0.59)	(0.58)	(1.01)	(2.21)	(0.75)	(0.83)	(0.05)	(0.43)	(1.37)
HML	−0.19***	−0.24***	−0.25***	−0.22**					−0.01	−0.16*	−0.04	−0.05	−0.04	−0.06	−0.09
	(−3.37)	(−4.68)	(−2.83)	(−2.13)					(−0.05)	(−1.79)	(−0.35)	(−0.48)	(−0.33)	(−0.51)	(−0.83)
RMW				−0.08					−0.26***	−0.36***	−0.08	−0.12*	−0.07	−0.12	−0.21**
				(−0.94)					(−2.60)	(−3.97)	(−1.08)	(−1.68)	(−0.71)	(−1.38)	(−2.38)
CMA				−0.02					−0.12	−0.20*	−0.06	−0.09	−0.07	−0.08	−0.14
				(−0.21)					(−1.02)	(−1.65)	(−0.54)	(−0.74)	(−0.64)	(−0.67)	(−1.19)
UMD		−0.08**	−0.07	−0.07		−0.05									
		(−2.12)	(−1.48)	(−1.43)		(−0.79)									
STR		0.33***	0.40***	0.40***		0.39***									
		(5.13)	(4.36)	(4.36)		(4.33)									
LTR		0.04	0.17**	0.16**		0.14**									
		(0.59)	(2.17)	(2.01)		(2.11)									
LIQ			0.02	0.02											
			(0.51)	(0.49)											
R_{ME}					0.03	−0.02									
					(0.49)	(−0.38)									
R_{IA}					−0.18**	−0.28***									
					(−2.04)	(−2.68)									
R_{ROE}					−0.12*	0.00									
					(−1.68)	(0.04)									
R_{EG}						0.04									
						(0.43)									
MGMT							−0.08								
							(−0.83)								
PERF							−0.07								
							(−1.02)								
FIN								−0.09							
								(−1.46)							
PEAD								−0.20*							
								(−1.89)							
X									0.23**	−0.29***	−0.08	−0.03	−0.07	−0.01	−0.13
									(2.07)	(−4.27)	(−1.37)	(−0.46)	(−0.85)	(−0.07)	(−1.56)
T	1086	1086	660	660	672	672	648	606	714	714	714	714	714	702	714
R^2	0.09	0.21	0.25	0.25	0.11	0.24	0.11	0.10	0.11	0.15	0.11	0.10	0.10	0.10	0.11

Table A.9:

Portfolio Analysis and Limit to Arbitrage

The table reports the abnormal returns (alphas) of the low-PS, high-PS and the directional high minus low signal portfolios (D-HML) obtained as difference between high-PS and low-PS for the low (bottom 33%) and high (top 33%) limit to arbitrage proxies. The table reports results based on Fama and French (1993) 3-factor model (FF3) and the Fama and French (2015) 5-factor model (FF5). Limit to arbitrage is measured as: i) firm size measured with the market capitalization; ii) Amihud (2002) illiquidity; iii) idiosyncratic volatility estimated on the residuals after the Fama and French (1993) 3-factor model on daily returns over the past 2 months. The sample period is from July 1932 to December 2022. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

		Low PS			High PS			D-HML		
		(1)	(2)	Diff.	(1)	(2)	Diff.	(1)	(2)	Diff.
FF3	Size: Big (1) vs. Small (2)	−0.39***	−1.67	−1.28***	0.38***	0.94	0.56***	0.77***	2.61	1.84***
		(−4.50)		(−7.36)	(5.71)		(3.47)	(5.88)		(6.92)
	Amihud Illiq.: Low (1) vs. High (2)	−0.31***	−1.62	−1.31***	0.36***	0.72	0.36***	0.67***	2.33	1.66***
		(−3.88)		(−7.07)	(6.14)		(2.64)	(5.32)		(6.40)
	Idio. Vol.: Low (1) vs. High (2)	−0.23***	−1.81	−1.58***	0.47***	−0.05	−0.52***	0.70***	1.76	1.07***
		(−2.80)		(−8.10)	(6.54)		(−2.70)	(5.25)		(3.67)
FF5	Size: Big (1) vs. Small (2)	−0.45***	−1.80	−1.35***	0.52***	1.18	0.66***	0.97***	2.98	2.01***
		(−4.14)		(−6.17)	(6.58)		(3.42)	(5.35)		(6.11)
	Amihud Illiq.: Low (1) vs. High (2)	−0.43***	−1.79	−1.36***	0.52***	0.87	0.36**	0.95***	2.66	1.71***
		(−3.67)		(−5.76)	(6.60)		(2.32)	(5.17)		(5.60)
	Idio. Vol.: Low (1) vs. High (2)	−0.39***	−1.67	−1.28***	0.53***	0.42	−0.11	0.92***	2.09	1.17***
		(−3.68)		(−6.52)	(5.63)		(−0.57)	(5.01)		(3.58)

Table A.10:

Directional Signal, Momentum and Directional Momentum - Stock with Traded Options

Each month, value-weighted decile portfolios are sorted according to the directional signal PS (Panel A) and the momentum MOM (Panel B). The table reports raw excess and abnormal returns (alphas) of portfolios obtained as difference between the portfolio built using stocks in the top decile minus the portfolio in the bottom decile for the subset of stocks with traded options. Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); (v) the Hou et al. (2020a) 5-factor model (Q5); v) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurstosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtotis of the portfolio returns, respectively. The sample period is from January 1996 to December 2020. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Panel A. Directional signal high minus low (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.95***	0.91***	0.92***	0.95***	0.93***	1.04***	0.96***
	(3.67)	(3.61)	(3.39)	(3.51)	(3.06)	(3.77)	(3.11)
MKT		0.06	0.05	0.04	−0.04	0.01	−0.05
		(0.76)	(0.63)	(0.58)	(−0.51)	(0.09)	(−0.64)
SMB			−0.01	−0.07	−0.11		
			(−0.11)	(−0.68)	(−1.15)		
HML			−0.18	−0.18	−0.27		
			(−1.40)	(−1.07)	(−1.54)		
RMW				−0.11	−0.10		
				(−0.75)	(−0.64)		
CMA				0.09	0.10		
				(0.48)	(0.51)		
UMD					−0.00		0.06
					(−0.06)		(0.62)
STR					0.26**		0.25**
					(2.44)		(2.17)
LTR					0.13		0.09
					(0.74)		(0.54)
R_{ME}						−0.05	−0.09
						(−0.54)	(−0.95)
R_{IA}						−0.16	−0.21
						(−1.09)	(−1.18)
R_{ROE}						−0.05	−0.12
						(−0.37)	(−0.74)
R_{EG}						−0.07	0.02
						(−0.40)	(0.12)
T	300	300	300	300	300	300	300
R^2	0.04	0.04	0.05	0.05	0.09	0.04	0.07
Vol	4.35	4.35	4.35	4.35	4.35	4.35	4.35
SR	0.62	0.62	0.62	0.62	0.62	0.62	0.62
Skew	0.50	0.50	0.50	0.50	0.50	0.50	0.50

Panel B. Momentum winners minus losers (WML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.48	1.02**	1.06**	0.63	−0.07	0.01	0.06
	(0.85)	(2.16)	(2.56)	(1.11)	(−0.24)	(0.02)	(0.16)
MKT		−0.76***	−0.84***	−0.68***	−0.04	−0.43*	−0.09
		(−3.11)	(−4.11)	(−3.79)	(−0.38)	(−1.74)	(−0.55)
SMB			0.17	0.36	0.22		
			(0.79)	(1.58)	(1.41)		
HML			−0.97***	−1.40***	−0.28		
			(−2.69)	(−3.60)	(−0.89)		
RMW				0.60	0.19		
				(1.65)	(0.81)		
CMA				0.51	0.13		
				(0.89)	(0.50)		
UMD					1.61***		1.64***
					(17.79)		(23.34)
STR					−0.40***		−0.41***
					(−3.70)		(−3.44)
LTR					−0.25		−0.30*
					(−1.49)		(−1.79)
R_{ME}						0.77**	0.15
						(2.44)	(1.17)
R_{IA}						−1.08**	−0.14
						(−2.19)	(−0.63)
R_{ROE}						1.48***	0.00
						(3.82)	(0.01)
R_{EG}						0.48	−0.00
						(1.24)	(−0.01)
T	300	300	300	300	300	300	300
R^2	0.00	0.10	0.18	0.20	0.71	0.27	0.71
Vol	10.84	10.84	10.84	10.84	10.84	10.84	10.84
SR	0.10	0.10	0.10	0.10	0.10	0.10	0.10
Skew	-0.93	-0.93	-0.93	-0.93	-0.93	-0.93	-0.93

Table A.11:

Accuracy Analysis

This table reports on the predictive ability of the probability forecasts (PS) and the momentum (MOM) strategies. Every month we regress 1-month ahead $(t+1)$ return percentiles (left table) and returns (right table) against a series of dummy variables that are equal to 1 if a stock was classified in month t in the indicated decile following a specific strategy and zero otherwise. The constant term represents the 5-th decile so that each coefficient indicates the changes in the dependent variable as we move from decile 5 to the indicated decile. The estimation is rolled over using the Fama and MacBeth (1973) methodology. Newey and West (1987) corrected t-stats are in parentheses. *, ** and *** indicate a 10%, 5% and 1% significance level, respectively.

	Percentiles		Returns	
	PS	MOM	PS	MOM
Constant (Decile 5)	50.03***	50.93***	1.40***	1.37***
	(566.50)	(361.94)	(5.57)	(6.83)
Decile 1	−3.06***	−4.68***	−0.93***	−0.28
	(−18.75)	(−13.80)	(−12.69)	(−1.47)
Decile 2	−1.63***	−1.87***	−0.54***	−0.09
	(−11.37)	(−8.84)	(−9.46)	(−0.90)
Decile 3	−0.89***	−1.06***	−0.26***	−0.09
	(−7.33)	(−7.17)	(−5.00)	(−1.48)
Decile 4	−0.39***	−0.33***	−0.11**	0.02
	(−3.00)	(−3.07)	(−2.20)	(0.35)
Decile 6	0.41***	0.32***	0.13**	0.07*
	(3.01)	(2.99)	(2.00)	(1.81)
Decile 7	0.65***	0.58***	0.14**	0.11**
	(4.83)	(4.66)	(2.18)	(2.11)
Decile 8	1.20***	0.69***	0.34***	0.20***
	(8.34)	(4.41)	(5.28)	(2.95)
Decile 9	1.59***	0.64***	0.48***	0.29***
	(11.57)	(3.35)	(6.74)	(3.78)
Decile 10	2.15***	0.14	0.78***	0.48***
	(12.97)	(0.52)	(9.61)	(3.98)

Table A.12:

Value-Weighted Univariate Portfolio Analysis With Transaction Costs

Each month, value-weighted decile portfolios are sorted according to the directional signal PS (Panel A) and the momentum MOM (Panel B). The table reports raw excess and abnormal returns (alphas) of portfolios obtained as difference between the portfolio built using stocks in the top decile minus the portfolio in the bottom decile. Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); (v) the Hou et al. (2020a) 5-factor model (Q5); v) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurstosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtotis of the portfolio returns, respectively. The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5 model. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Panel A. Directional signal high minus low (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.42^{***}	0.42^{***}	0.48^{***}	0.86^{***}	0.75^{***}	1.04^{***}	0.82^{***}
	(3.52)	(3.84)	(4.71)	(6.77)	(5.06)	(6.30)	(4.95)
MKT		0	0.02	0.06	−0.03	0.04	−0.03
		(−0.06)	(0.49)	(1.52)	(−0.94)	(0.84)	(−0.89)
SMB			0.08	0.06	−0.01		
			(1.62)	(1.04)	(−0.29)		
HML			-0.22^{***}	−0.1	-0.25^{***}		
			(−4.59)	(−1.00)	(−2.77)		
RMW				−0.11	−0.05		
				(−1.48)	(−0.60)		
CMA				−0.05	−0.01		
				(−0.41)	(−0.12)		
UMD					−0.05		−0.02
					(−1.18)		(−0.34)
STR					0.41^{***}		0.4^{***}
					(5.45)		(5.53)
LTR					0.17^{**}		0.15^{**}
					(2.34)		(2.19)
R_{ME}						0.03	−0.02
						(0.51)	(−0.38)
R_{IA}						-0.19^{**}	-0.3^{***}
						(−2.17)	(−3.10)
R_{ROE}						−0.08	0.01
						(−1.07)	(0.10)
R_{EG}						−0.13	−0.04
						(−1.39)	(−0.43)
T	1086	1086	1086	714	714	672	672
R^2	0.01	0.01	0.07	0.09	0.24	0.10	0.22
Vol	3.35	3.35	3.35	3.33	3.33	3.39	3.39
SR	0.16	0.16	0.16	0.49	0.49	0.50	0.50
Skew	0.22	0.22	0.22	0.72	0.72	0.69	0.69
Kurtosis	8.63	8.63	8.63	6.22	6.22	6.06	6.06

Panel B. Momentum winners minus losers (WML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.66^{***}	1.09^{***}	1.29^{***}	1.15^{***}	0.13	−0.19	−0.05
	(2.66)	(6.24)	(6.79)	(3.62)	(1.08)	(−0.60)	(−0.34)
MKT		-0.54^{***}	-0.39^{***}	-0.34^{***}	-0.08^{**}	-0.18^{*}	−0.06
		(−4.84)	(−4.84)	(−3.28)	(−2.04)	(−1.89)	(−1.53)
SMB			−0.22	−0.02	−0.06		
			(−1.64)	(−0.13)	(−0.98)		
HML			-0.7^{***}	-0.81^{***}	−0.01		
			(−5.36)	(−3.52)	(−0.13)		
RMW				0.34	0.06		
				(1.08)	(0.80)		
CMA				0.5	−0.02		
				(1.49)	(−0.22)		
UMD					1.48^{***}		1.43^{***}
					(47.28)		(38.90)
STR					-0.14^{**}		-0.14^{**}
					(−2.36)		(−2.46)
LTR					−0.04		−0.01
					(−0.57)		(−0.10)
R_{ME}						0.35^{*}	−0.01
						(1.68)	(−0.15)
R_{IA}						−0.26	−0.11
						(−1.09)	(−1.33)
R_{ROE}						1.15^{***}	0.09
						(4.74)	(1.33)
R_{EG}						0.76^{***}	0.21^{**}
						(3.29)	(1.99)
T	1086	1086	1086	714	714	672	672
R^2	0.01	0.13	0.24	0.13	0.83	0.31	0.83
Vol	7.78	7.78	7.78	7.13	7.13	7.27	7.27
SR	0.17	0.17	0.17	0.29	0.29	0.26	0.26
Skew	-2.70	-2.70	-2.70	-1.37	-1.37	-1.37	-1.37
Kurtosis	22.65	22.65	22.65	9.70	9.70	9.48	9.48

Table A.13:

Value-Weighted Univariate Portfolio Analysis on Directional Signal (PS) and Momentum (MOM) Excluding Stocks Below 5$

Each month, value-weighted decile portfolios are sorted according to the directional signal PS (Panel A) and according to the momentum MOM (Panel B). We remove stocks for which the previous monthly closing price was below 5$. The table reports raw excess and abnormal returns (alphas) of portfolios obtained as difference between the portfolio built using stocks in the top decile minus the portfolio built using the bottom decile of the signal and momentum. Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); (v) the Hou et al. (2020a) 5-factor model (Q5); v) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurstosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtotis of the portfolio returns, respectively. The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5 model. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Panel A. Directional signal high minus low (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.74***	0.72***	0.76***	0.98***	0.87***	1.08***	0.87***
	(7.18)	(6.74)	(6.59)	(6.97)	(4.79)	(6.17)	(4.50)
MKT		0.03	0.02	0.04	−0.05	0.02	−0.04
		(0.82)	(0.90)	(0.91)	(−1.33)	(0.53)	(−0.98)
SMB			0.11***	0.05	−0.02		
			(2.71)	(0.87)	(−0.31)		
HML			−0.15***	−0.05	−0.19**		
			(−3.11)	(−0.50)	(−2.06)		
RMW				−0.15*	−0.09		
				(−1.95)	(−1.13)		
CMA				−0.09	−0.06		
				(−0.78)	(−0.56)		
UMD					−0.04		−0.03
					(−0.85)		(−0.48)
STR					0.38***		0.37***
					(4.10)		(3.91)
LTR					0.16**		0.16**
					(2.22)		(2.41)
R_{ME}						0.06	0
						(0.89)	(0.07)
R_{IA}						−0.19**	−0.31***
						(−2.29)	(−2.84)
R_{ROE}						−0.09	0
						(−1.13)	(0.04)
R_{EG}						−0.04	0.05
						(−0.45)	(0.58)
T	1086	1086	1086	714	714	672	672
R^2	0.06	0.06	0.09	0.10	0.23	0.11	0.22
Vol	3.06	3.06	3.06	3.23	3.23	3.29	3.29
SR	0.54	0.54	0.54	0.61	0.61	0.62	0.62
Skew	0.87	0.87	0.87	0.65	0.65	0.62	0.62
Kurtosis	7.91	7.91	7.91	5.98	5.98	5.84	5.84

Panel B. Momentum winners minus losers (WML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.97***	1.29***	1.46***	1.41***	0.42***	0.12	0.22
	(4.85)	(8.02)	(8.81)	(4.81)	(3.47)	(0.40)	(1.59)
MKT		−0.41***	−0.31***	−0.26***	−0.02	−0.1	0.01
		(−5.09)	(−4.39)	(−2.98)	(−0.67)	(−1.25)	(0.33)
SMB			−0.04	0.04	−0.01		
			(−0.27)	(0.20)	(−0.25)		
HML			−0.62***	−0.75***	−0.03		
			(−5.23)	(−3.88)	(−0.48)		
RMW				0.17	−0.08		
				(0.57)	(−1.38)		
CMA				0.44	−0.07		
				(1.53)	(−0.81)		
UMD					1.4***		1.38***
					(30.49)		(32.68)
STR					−0.08*		−0.08*
					(−1.66)		(−1.89)
LTR					0.02		0.01
					(0.32)		(0.12)
R_{ME}						0.4*	0.05
						(1.92)	(0.98)
R_{IA}						−0.22	−0.1
						(−1.03)	(−1.40)
R_{ROE}						0.92***	−0.08
						(4.58)	(−1.21)
R_{EG}						0.78***	0.27***
						(3.57)	(2.60)
T	1086	1086	1086	714	714	672	672
R^2	0.02	0.11	0.20	0.12	0.85	0.28	0.85
Vol	6.89	6.89	6.89	6.47	6.47	6.57	6.57
SR	0.35	0.35	0.35	0.46	0.46	0.43	0.43
Skew	-2.39	-2.39	-2.39	-0.78	-0.78	-0.79	-0.79
Kurtosis	23.51	23.51	23.51	7.56	7.56	7.46	7.46

Table A.14:

Equally-Weighted Univariate Portfolio Analysis on Directional Signal (PS) and Momentum (MOM)

Each month, equally-weighted decile portfolios are sorted according to the directional signal PS (Panel A) and according to the momentum MOM (Panel B). The table reports raw excess and abnormal returns (alphas) of portfolios obtained as difference between the portfolio built using stocks in the top decile minus the portfolio built using the bottom decile of the signal and momentum. Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Mossin (1966) with MKT factor; ii) the three-factor model of Fama and French (1993) with MKT, SMB, and HML factors (FF3); iii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iv) the FF5 model augmented by the momentum, short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); v) the Hou et al. (2020a) 5-factor model (Q5); vi) the Q5 model augmented by the momentum, short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). T is the number of months in the sample period. Vol, SR, Skew and Kurstosis are the standard deviation, annualized Sharpe ratio, Skewness and Kurtotis of the portfolio returns, respectively. The sample runs from July 1932 to December 2022 for the CAPM and FF3, July 1963 to December 2022 for the FF5 and January 1967 to December 2022 for the Q5 model. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance level of 10%, 5% and 1%, respectively.

Panel A. Directional signal high minus low (D-HML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	1.70***	1.69***	1.76***	1.96***	1.85***	2.09***	1.83***
	(14.19)	(15.44)	(15.30)	(14.43)	(15.57)	(12.26)	(13.16)
MKT		0.01	0.09***	0.15***	0.05*	0.13***	0.06*
		(0.20)	(2.77)	(4.17)	(1.96)	(3.40)	(1.96)
SMB			−0.20*	0.06	0.01		
			(−1.77)	(0.83)	(0.22)		
HML			−0.20***	0.04	−0.09		
			(−2.66)	(0.51)	(−1.33)		
RMW				0.10	0.13**		
				(1.37)	(2.30)		
CMA				−0.01	0.10		
				(−0.11)	(0.99)		
UMD					−0.06*		−0.08*
					(−1.79)		(−1.86)
STR					0.44***		0.44***
					(8.24)		(8.17)
LTR					0.04		0.03
					(0.73)		(0.56)
R_{ME}						0.04	0.03
						(0.47)	(0.59)
R_{IA}						0.04	0.01
						(0.58)	(0.13)
R_{ROE}						0.01	0.12*
						(0.22)	(1.80)
R_{EG}						−0.10	0.03
						(−1.10)	(0.30)
T	1086	1086	1086	714	714	672	672
R^2	0.19	0.19	0.25	0.35	0.49	0.35	0.49
Vol	3.51	3.51	3.51	3.03	3.03	3.09	3.09
SR	1.42	1.42	1.42	1.98	1.98	1.97	1.97
Skew	-1.86	-1.86	-1.86	1.01	1.01	0.99	0.99
Kurtosis	28.02	28.02	28.02	6.40	6.40	6.26	6.26

Panel B. Momentum winners minus losers (WML)

	Raw return	CAPM	FF3	FF5	FF5+UMD +S/L-TR	Q5	Q5+UMD +S/L-TR
Alpha	0.86***	1.16***	1.37***	0.96***	0.14	−0.07	0.03
	(3.84)	(6.77)	(7.87)	(3.34)	(0.89)	(−0.22)	(0.15)
MKT		−0.45***	−0.25***	−0.11	0.10*	0.01	0.09*
		(−2.95)	(−2.98)	(−1.25)	(1.82)	(0.07)	(1.67)
SMB			−0.50***	−0.22	−0.23***		
			(−3.78)	(−1.47)	(−3.74)		
HML			−0.70***	−0.49**	0.16		
			(−3.19)	(−2.56)	(1.60)		
RMW				0.49*	0.26***		
				(1.81)	(3.54)		
CMA				0.48*	0.10		
				(1.66)	(0.76)		
UMD					1.19***		1.05***
					(18.57)		(18.90)
STR					−0.09		−0.09
					(−1.05)		(−1.05)
LTR					−0.10		−0.00
					(−1.21)		(−0.04)
R_{ME}						0.10	−0.16***
						(0.57)	(−2.92)
R_{IA}						0.14	0.24*
						(0.58)	(1.80)
R_{ROE}						1.25***	0.48***
						(5.66)	(5.31)
R_{EG}						0.35*	−0.05
						(1.95)	(−0.49)
T	1146	1146	1146	714	714	672	672
R^2	0.01	0.11	0.25	0.11	0.69	0.33	0.70
Vol	7.63	7.63	7.63	6.20	6.20	6.35	6.35
SR	0.27	0.27	0.27	0.35	0.35	0.32	0.32
Skew	-4.24	-4.24	-4.24	-2.56	-2.56	-2.52	-2.52
Kurtosis	42.38	42.38	42.38	20.77	20.77	20.01	20.01

Table A.15:

High (minus) Low Value-Weighted Portfolios: Alternative Estimations

The Table reports abnormal returns (alphas) for the high (minus) low probability score portfolios. Abnormal returns (alphas) are generated based on different sets of asset pricing models: i) the capital asset pricing model (CAPM); ii) the five-factor model of Fama and French (2015) with MKT, SMB, HML, RMW, and CMA factors (FF5); iii) the FF5 model augmented by momentum, the short-term and long-term reversal momentum factors (FF5+UMD+S/L-TR); iv) the 5-factor (Q5) Hou et al. (2020a); v) the Q5 model augmented by momentum, the short-term and long-term reversal momentum factors (Q5+UMD+S/L-TR). The probability score is estimated using various specifications. Baseline refers to the probability score estimated using durations, lagged idiosyncratic volatility (IV) and lagged market return (Mkt). CKX refers to a model in which the Conrad et al. (2014) variables are used while Kelly refers to a model in which the Kelly variables are used. Vol, SR, Skew and Kurt indicate the volatility, annualized Sharpe ratio, skewness and kurtosis of each portfolio, respectively. Newey-West adjusted t-statistics are given in parentheses. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

	Baseline	Duration Only	Duration + IV	Duration + Mkt	Mkt+IV	Baseline + CKX	CKX Only	Baseline + Kelly	Kelly Only
CAPM	0.76***	0.58***	0.67***	0.74***	1.51***	0.91***	0.14	2.02***	1.89***
	(6.86)	(6.27)	(6.51)	(7.30)	(6.01)	(6.10)	(0.74)	(9.06)	(7.33)
FF5	1.02***	0.77***	0.86***	1.05***	1.02***	0.78***	−0.11	1.59***	1.32***
	(7.62)	(4.52)	(5.51)	(7.14)	(5.41)	(5.75)	(-0.73)	(8.61)	(6.54)
FF5+UMD+S(L)TR	0.92***	0.80***	0.89***	1.00***	0.81***	0.61***	−0.18	1.16***	0.94***
	(5.28)	(4.03)	(4.77)	(5.56)	(3.28)	(4.23)	(-1.07)	(5.67)	(3.92)
Q5	1.16***	0.66***	0.75***	1.23***	0.62**	0.67***	−0.33*	1.09***	0.77***
	(6.41)	(3.73)	(4.22)	(6.97)	(2.16)	(3.97)	(-1.84)	(5.37)	(3.32)
Q5+UMD+S(L)TR	0.93***	0.73***	0.82***	1.03***	0.61**	0.51***	−0.25	0.85***	0.67***
	(4.90)	(3.31)	(3.99)	(5.38)	(2.01)	(3.20)	(-1.34)	(4.00)	(2.67)
Vol	0.03	0.03	0.03	0.03	0.09	0.04	0.06	0.06	0.07
SR	0.54	0.38	0.43	0.58	0.23	0.31	-0.3	0.88	0.58
Skew	1.02	1.42	1.16	1.22	-0.98	-0.77	-1.34	-0.21	-0.52
Kurt	10.52	13.93	10.47	8.73	11.55	9.58	11.37	8.6	9.02

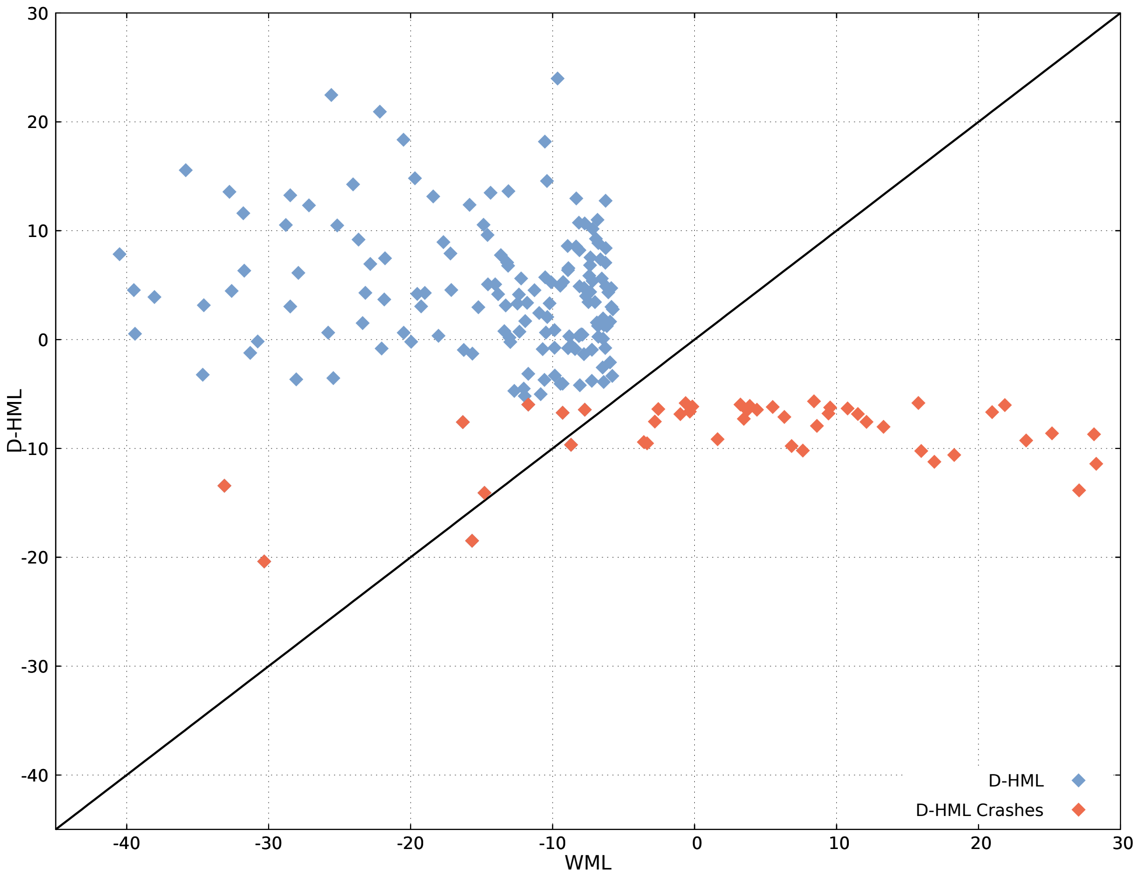


Figure A.1: This figure contains the scatter of the 3-month cumulative returns of the D-HML against the WML strategy conditional on either one of the strategies having returns smaller than the 5th percentile of the D-HML strategy. Orange points represent occasions where D-HML was below its 5th percentile, whereas blue points represent occasions where WML was below D-HML's 5th percentile, but D-HML was not.

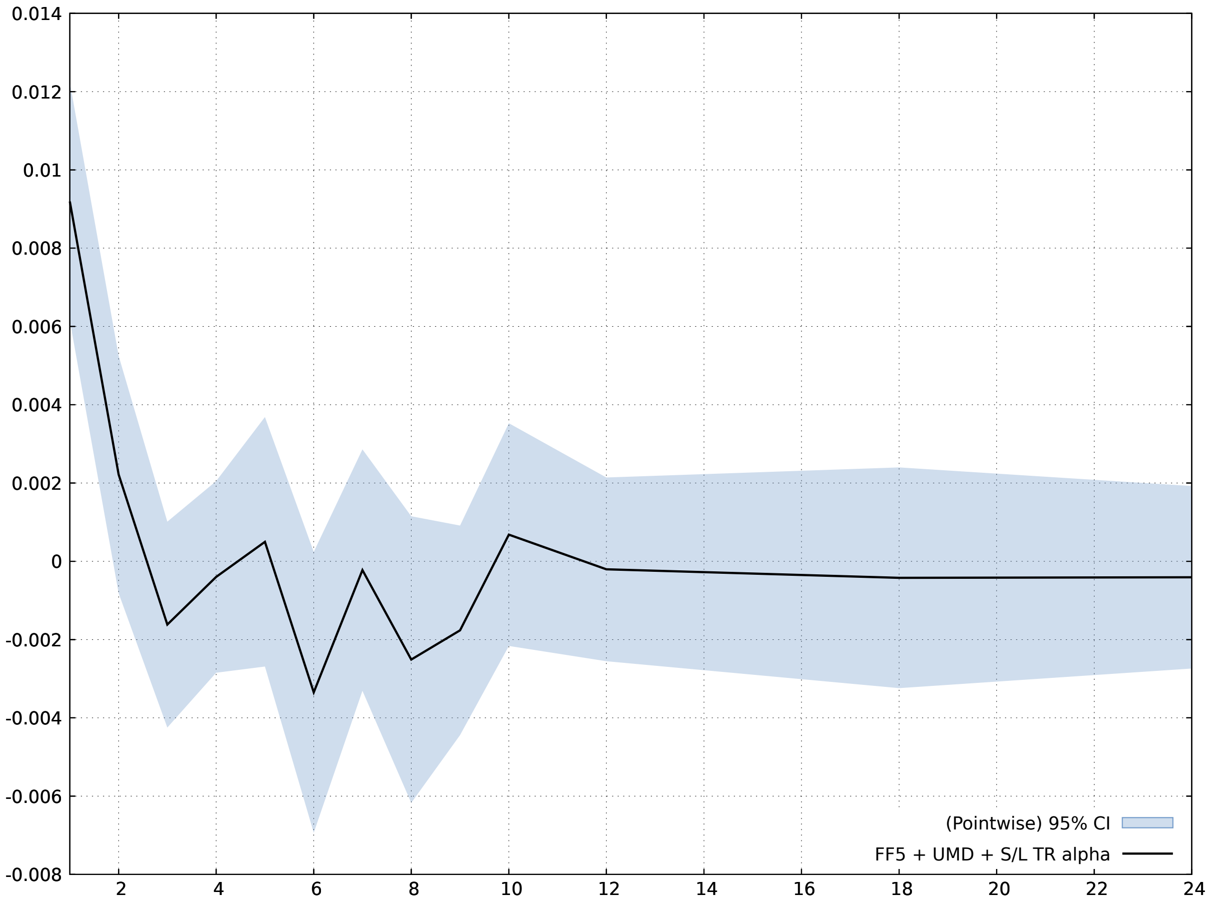


Figure A.2: This figure contains the alpha of the D-HML portfolio estimated from the FF5 + UMD + S/L TR model for holding periods from 1 to 24 months after the portfolio formation.

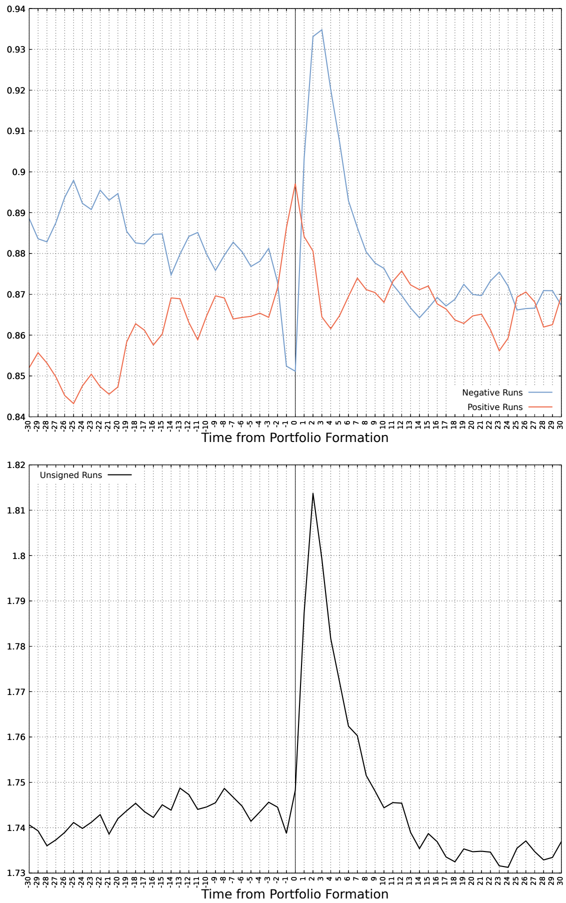


Figure A.3: This figure contains in the left plot the average length (run) of consecutive days with negative and positive returns around earnings announcements. The right plot contains the average length of runs regardless of their direction.

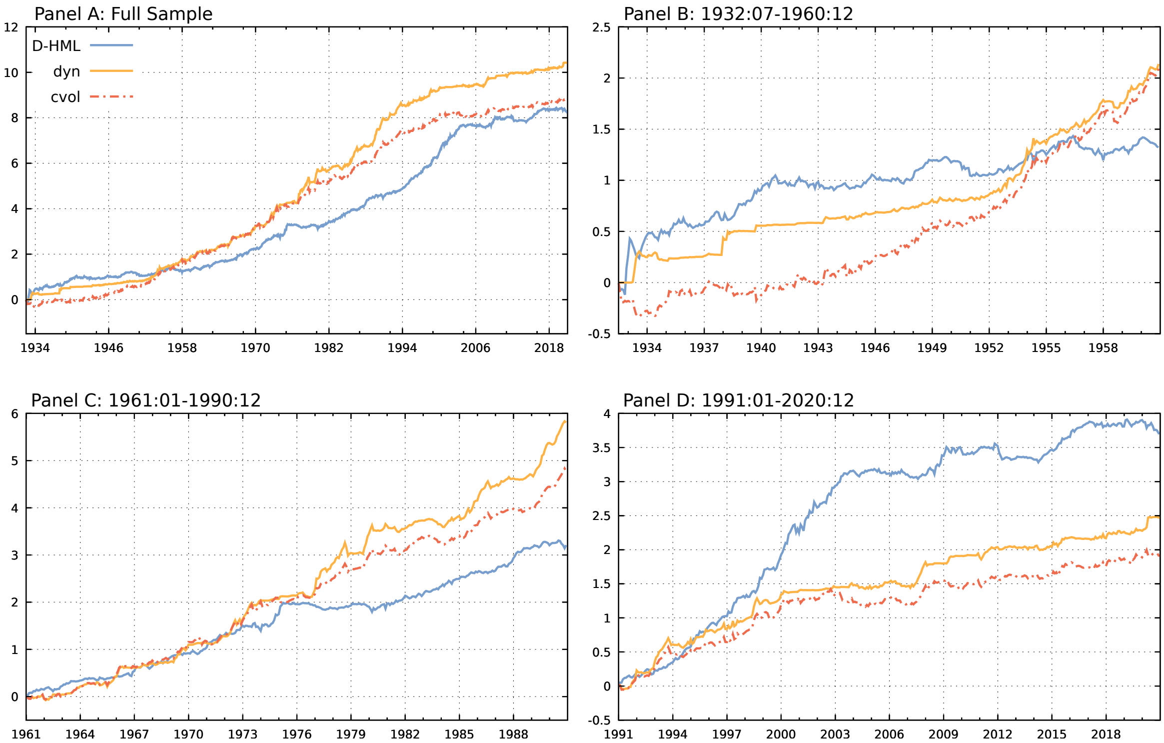


Figure A.4: This figure contains the cumulative log return of the D-HML (blue continuous line), the Daniel and Moskowitz's 2016 dynamic (dyn) WML (yellow continuous line) and the Barroso and Santa-Clara's 2015 constant volatility (cvol) WML (red dashed line) strategies. Panel A plots the cumulative returns over the full sample period from 1932:07 to 2020:12. Panels B–D plot the returns over three roughly 30 years sub-samples: 1932–1960, 1961–1990 and 1991–2020.