All Days Are Not Created Equal:
Understanding Momentum by Learning To Weight Past Returns *

Heiner Beckmeyer† Timo Wiedemann‡

This version: September 9, 2024

Abstract

By flexibly weighting the information contained in past realized returns, we construct a momentum strategy that not only outperforms and subsumes the performance of traditional stock momentum, performs well in crises and continues to work in the most recent decades, but that also allows us to study the origins of momentum profits. We investigate how past returns are optimally weighted to extract momentum signals and find that the response to earnings announcements, market-wide jumps and the largest individual returns realized in the formation period are most informative about future stock returns. We show that the way past returns are weighted is in line with the strategy exploiting an underreaction to the information contained in past realized returns, and rule out alternative behavioral and risk-based explanations.

Keywords: Momentum, Machine Learning, Big Data, Anomalies

JEL classification: G11, G12, G17

*We want to thank participants at the 16th Annual Conference of the Behavioural Finance Working Group and the 2023 Annual Meeting of the German Finance Association (DGF), as well as seminar participants at the University of Münster for valuable feedback and discussions.

†University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany. Tel: +49 251 83 22771. E-mail: heiner.beckmeyer@wiwi.uni-muenster.de.

‡University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany. Tel: +49 251 83 22885. E-mail: timo.wiedemann@wiwi.uni-muenster.de.

1. Introduction

Momentum is one of the most prominent return anomalies and has garnered considerable interest by academics and practitioners. The two seminal studies by Jegadeesh and Titman (1993) and Carhart (1997) have not only sparked a large and growing body of academic studies, but have also prompted practitioners to translate the academic evidence on return continuation into numerous trading vehicles: many hedge funds, commodity trading advisors, and mutual funds employ momentum-based strategies. Today, momentum ETFs are a low-cost alternative for investors wishing to obtain exposure to trend-following investment styles.

Despite its practical and academic success, there is still no consensus about the origins of momentum profits. Prominent explanations include a misreaction to news but both an underreaction and a delayed overreaction have been proposed. For example, in the model of Daniel, Hirshleifer, and Subrahmanyam (1998), investors are overconfident in their own signals, leading to a delayed overreaction to news, whereas investors in the model by Barberis, Shleifer, and Vishny (1998) are conservative, such that they underreact to new information. Hong and Stein (1999) propose a combination of the two theories in a single framework. As an alternative explanation, Kelly, Moskowitz, and Pruitt (2021) show empirically that momentum profits are well-explained by time-variation in the compensation for risk in the stock market.

In this paper, we propose a simple change to how momentum signals are generated from past realized returns. Consistent with the adage that *all days are not created equal*, we allow the weights placed on returns realized in the formation period to differ from an equal importance assumed by standard momentum. We do so by setting up a simple machine learning framework that extracts these weights from the formation period returns themselves and firm characteristics that serve as additional instruments. Importantly, we maintain the simple and well motivated structure of a *weighted* average of past realized returns as our momentum signal. We coin the resulting momentum strategy *characteristic-managed momentum (CMM)*.

We start by showing that changing the way past returns are aggregated to expected return signals significantly improves the performance of momentum strategies. Not only is CMM vastly more profitable, but it is also significantly more robust. Unlike traditional momentum, CMM continues to work since 2003 and does not suffer from momentum crashes discussed by Daniel and Moskowitz (2016). Keep in mind that the sole difference

between CMM and the standard momentum strategy is in how the same daily returns realized in the formation period are weighted. We can show that CMM fully subsumes the performance of standard momentum. While the resulting value-weighted return of the CMM-based high-minus-low portfolio is significant within each standard momentum portfolio, we find no significant return spreads of sorting by the traditional momentum signal, once we control for the optimized CMM signal.

To understand why CMM is better able to extrapolate from past returns, we first show that the average weights placed on past returns are sparse: on average, just two of the 231 returns in the formation period receive a weight of 30% – a 34-fold increase compared to standard momentum's equal weighting. Just 30 returns already receive more than half of the total weight, suggesting that some returns are significantly more informative than postulated by equal weights, and others significantly less informative. To understand which days are most informative about return extrapolation, we investigate if the weights are systematically elevated on certain news and event days. We find that days with particularly large returns of the stock in question are most important, followed by days with market-wide jumps. We also find that the return response to earnings and FOMC announcements are overweighted on average. Importantly, we show that solely basing the momentum signals on the return reactions to these news and event days is insufficient to generate profitable signals that exceed the performance of standard momentum. Instead, CMM's conditional weighting function is required to adequately extract information about how returns will continue in the future.

We next seek to understand what is driving momentum profits in the stock market, using our optimized CMM strategy as the benchmark. We differentiate between behavioral and risk-based explanations. For the former, we devise a simple empirical test allowing us to differentiate an over- from an underreaction to information released in the formation period by investigating the correlation between absolute returns and weights. We find that CMM predominantly extracts momentum signals from an underreaction to news. In fact, only among this subset of stocks, which captures between $60-90\%$ of the cross-section over time, do we find a statistically significant and economically meaningful value-weighted high-minus-low return spread. We corroborate this evidence of an underreaction to news causing momentum profits by showing that the information content of the CMM signal gradually declines, with the high-minus-low return spread converging to zero, with no long-term correction.

Finally, to rule out a risk-based explanation, we apply the methodology of Kelly et al.

(2021). We show that, unlike standard momentum, CMM is not explained by the time variation in the compensation for risk in the stock market. Instead, the compensation for risk and CMM's expected return signal are largely orthogonal and both provide information about the variation in future stock returns. Furthermore, CMM is not subject to momentum crashes. We find that the strategy specifically excels in identifying those momentum losers that outperform in the future, which avoids much of the option-like behavior of the standard momentum strategy discussed by Daniel and Moskowitz (2016).

We run several robustness checks and show that the discreteness of how information is released over the formation period cannot explain CMM's success (Da, Gurun, and Warachka, 2014), nor can it be explained by its exposure to common factors. Finally, we show that CMM outperforms alternative augmentations of momentum strategies that have been proposed in the literature.

Literature Review. This paper is part of a burgeoning literature on momentum strategies, following the seminal studies by Jegadeesh and Titman (1993) and Carhart (1997). Momentum has been shown to work in many asset classes (Menkhoff, Sarno, Schmeling, and Schrimpf, 2012; Asness, Moskowitz, and Pedersen, 2013; Heston, Jones, Khorram, Li, and Mo, 2023; Li, Yuan, and Zhou, 2023), exists in risk factors (Ehsani and Linnainmaa, 2022) and in returns of varying frequencies (Lou, Polk, and Skouras, 2019; Barardehi, Bogousslavsky, and Muravyev, 2022). However, in recent years momentum stopped working. Green, Hand, and Zhang (2017) show that momentum – along with many other characteristic-sorted investment strategies – ceased to generate significant returns after 2003. Furthermore, Daniel and Moskowitz (2016) show that momentum is prone to severe crashes. We propose to change how past returns are weighted to generate momentum investment signals. This seemingly simple change delivers a momentum strategy that shows great performance also after 2003 and is not subject to crashes, showcasing that past realized returns on some days are more informative than on others.

There is still an active debate about the origin of momentum profits. Prominent explanations include an under- or an overreaction to news (Daniel et al., 1998; Barberis et al., 1998). For example, in the model of Hong and Stein (1999), two types of agents consume different streams of information. One investor type impounds private news, which diffuse only gradually among them, into market prices, leading to an initial underreaction. The other investor type follows the recent trend, which leads to a delayed overreaction. In contrast, Kelly et al. (2021) favor a risk-based explanation and show that momentum

signals are a proxy for time variation in the compensation for risk in the stock market. By investigating the weights our momentum strategy places on returns realized in the formation period, we find compelling evidence that an underreaction to news drives its success. In adjacent empirical tests, we rule out alternative (risk-based) explanations.

The algorithm proposed is part of a recent shift towards using machine learning methods to predict future returns in various asset classes (Gu, Kelly, and Xiu, 2020; Chen, Pelger, and Zhu, 2023; Bali, Goyal, Huang, Jiang, and Wen, 2020; Bali, Beckmeyer, Moerke, and Weigert, 2023).

Finally, our paper is related to the literature identifying important days for explaining or predicting stock returns. To this end, Savor and Wilson (2013) and Ernst, Gilbert, and Hrdlicka (2019) show that market returns are significantly higher on days with macroeconomic news releases. Liang (2003) argues that earnings announcements provide particularly important information to investors and Savor and Wilson (2016) show that stock returns are amplified around earnings announcements. Baker, Bloom, Davis, and Sammon (2021) identify days with large stock-wide absolute returns and relate them to next-day newspaper articles to understand their drivers. Our machine learning learns to weight the market's reaction to news released throughout the formation period. We find that earnings announcements, and days with large market-wide and stock-specific price moves are particularly important for the generation of momentum signals.

2. Improving Momentum Signals

Traditional momentum in the spirit of Jegadeesh and Titman (1993) proposes that "winner" stocks, which have outperformed in the past, will continue to outperform "loser" stocks in the future. For this, the authors propose to compare the stocks' performance over the past year, excluding the most recent month to account for short-term return reversals. In essence, this approach assigns each of the 231 trading days of months $t-12m$ to $t-1m$ an equal importance to inform about the stock's return in month $t+1m$. The inherent return extrapolation proposed by standard momentum (SM),

$$\mathbb{E}_t^{\text{SM}}[r_{i,t+1m}] = \frac{1}{231}\sum_{d=22}^{252} r_{i,t-d}, \tag{1}$$

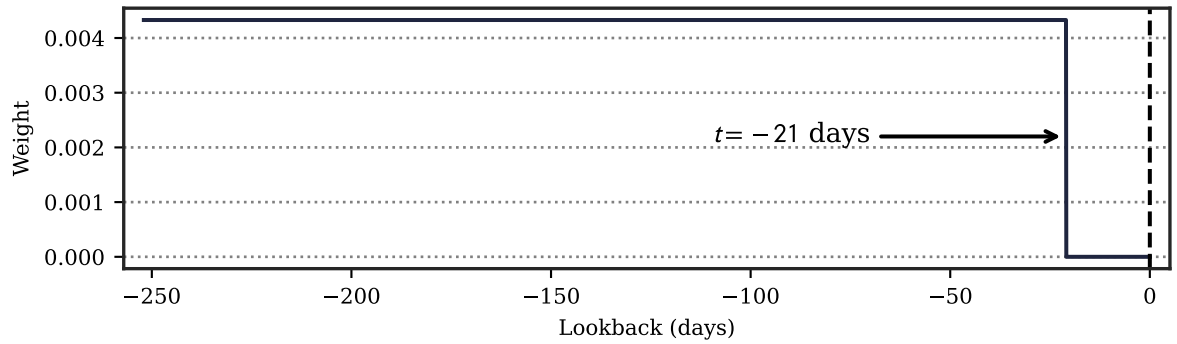


Fig. 1. Weights to Standard Momentum (SM)

The figure shows the weights assigned to past returns for the creation of the standard momentum (SM) signal. Over the past year, each of the 252 trading days is assigned an equal weight, excluding the 21 trading days of the most recent months to account for short-term return reversal effects.

is graphically depicted in Figure 1. Here, $r_{i,t}$ denotes the daily log return of stock i at time t and $r_{i,t+1m}$ the one-month ahead return of i. Because we are dealing with log returns, the sorting is the same regardless of whether we use the cumulative returns over the formation period, or the average return within the formation period.

While its simplicity is appealing, the implicit assumption that the return realized on each of the 231 trading days conveys equal information about future stock returns is highly restrictive and not supported by the literature. For example, Savor and Wilson (2013) and Ernst et al. (2019) show that market returns in general are significantly amplified on days with macroeconomic releases, Knox, Londono, Samadi, and Vissing-Jorgensen (2024) find that these same days carry significantly elevated equity premia, while Savor and Wilson (2016) and Engelberg, McLean, and Pontiff (2018) show that stock returns in general, and the returns of stock anomalies specifically, are higher around earnings announcements. Within the literature specifically targeting improvements to the momentum anomaly, we also find evidence that some days may be of higher importance: Novy-Marx (2012) shows that intermediate momentum – measured over months $t-12m$ to $t-7m$ – is more informative than using the return realized over the traditional formation period.

Following the overwhelming evidence from the literature that the return reaction on some days may be more important than that on other days, we propose a small change in how past returns are aggregated into momentum-based investment signals. Consistent with the adage that *all days are not created equal*, we estimate flexible weights w using modern tools from machine learning, instead of assigning a fixed weight to realized returns over the past year. Importantly, we maintain the economically-interpretable structure of

a linear *weighted* average of past returns:

$$\mathbb{E}_t^{\text{CMM}}[r_{i,t+1m}] = \sum_{d=22}^{252} w_{i,t-d} \cdot r_{i,t-d} \tag{2}$$

subject to boundary conditions,

$$\sum_{d=22}^{252} w_{i,t-d} = 1 \qquad \text{and} \qquad w_{i,t-d} \geq 0 \quad \forall\ d \in [22, 252]. \tag{3}$$

To see how and when the strategy's expected return signal can differ from that of standard momentum, it is instructive to rewrite Eq. (2) as:

$$\mathbb{E}_t^{\text{CMM}}[r_{i,t+1m}] = \underbrace{\sum_{d=22}^{252} (w_{i,t-d} - \overline{w}) \cdot r_{i,t-d}}_{\text{Deviation from SM}} + \underbrace{\sum_{d=22}^{252} \overline{w} \cdot r_{i,t-d}}_{\text{Momentum signal}} \quad \text{where} \quad \overline{w} = \frac{1}{231} \tag{4}$$

The signal is composed of two parts: the standard momentum signal, and a deviation thereof. If day d^+ is expected to be more important, it will receive a weight w_{t-d^+} greater than the average weight $\overline{w}$. Because of the boundary conditions, reflecting our desire to maintain the interpretable structure of a weighted average of past realized returns, this in turn requires underweighting another day d^-, relative to the equal weights assigned by standard momentum. In other words, the setup in Eq. (4) is forced to triage between which days are expected to be particularly important for the formation of momentum signals. If an equal importance of the days in the formation period is expected to generate the most information momentum signal, the signal in Eq. (4) will simply collapse to that of standard momentum. We make use of this circumstance in Section 4 to differentiate alternative explanations for why momentum in the stock market works.

Methodology. We use both time t observable firm characteristics, $\boldsymbol{z}_{i,t}$ and the daily returns realized over the formation period, $\boldsymbol{r}_{i,t-252:t-22}$, as firm i specific instruments to allow for variation in the weights placed on past returns.[1] This allows the model to assign different weights to past realized returns for two firms with the exact same firm characteristics, depending on their profile and hence informational content of past realized

[1]Using firm characteristics as instrument is frequently done in the empirical asset pricing literature, see Kelly, Pruitt, and Su (2019); Gu et al. (2020); Kelly et al. (2021); Bali et al. (2023).

returns:

$$\boldsymbol{w}_{i,t-252:t-22} = \phi(\boldsymbol{r}_{i,t-252:t-22}, \boldsymbol{z}_{i,t}). \tag{5}$$

Here, $\boldsymbol{w}_{i,t-252:t-22}$ denotes the vector of weights for the returns within the formation period. ϕ is therefore a function of observable firm characteristics $\boldsymbol{z}$ and realized daily returns $\boldsymbol{r}$ over the same formation period used for standard momentum. The sole difference to standard momentum is a departure from a fixed to a flexible weighting of past returns. The circumstance that we use firm characteristics to capture cross-sectional heterogeneity in the optimal degree of return extrapolation gives rise to the strategy's name, *Characteristic-Managed Momentum (CMM)*.

We estimate the weighting function ϕ, subject to the boundary conditions in Eq. (3), using tools from machine learning. With this, we join a growing strand of the financial literature, which uses machine learning techniques to improve the prediction of financial targets (Gu et al., 2020; Chen et al., 2023; van Binsbergen, Han, and Lopez-Lira, 2023). For each firm available at the end of month t, we collect vector $\boldsymbol{z}_{i,t}$ of size $C = 153$, comprised of standard accounting- and market-based characteristics of the firm (Jensen, Kelly, and Pedersen, 2023). We then extract a single nonlinear feature $\widehat{z}$ of the firm, which combines interactions between the firm's characteristics and nonlinear effects, by feeding $\boldsymbol{z}$ through a feed-forward neural network (FFN).[2] We follow Gu et al. (2020) for the architecture and use 32 neurons for the first, 16 for the second and 8 for the last hidden layer:

$$\widehat{z}_{i,t} = \text{FFN}(\boldsymbol{z}_{i,t}) \tag{6}$$

We then combine feature $\widehat{z}$, as the description of the firm's economic conditions, with the firm's realized returns over the past twelve months, $\boldsymbol{r}_{i,t-252:t-22}$. To isolate the impact of changing only the way momentum signals are generated from past returns, we exclude the most recent month in this formation period, as in standard momentum. The combination of returns and characteristics is done by calculating their dot-product similarity to come up with an importance Score of each of the 231 past daily returns which we denote by the vector $\text{Score}_{i,t-252:t-22}$:

$$\text{Score}_{i,t-252:t-22} = \widehat{z}_{i,t} \cdot \boldsymbol{r}_{i,t-252:t-22}. \tag{7}$$

[2]We choose the Gaussian error linear unit as the nonlinear activation function. Results using the more commonly used ReLU (Gu et al., 2020) are similar but slightly weaker.

Finally, we normalize the 231 SCOREs per firm×month observation to obtain the weight for the past daily return at $t-d$ as

$$w_{i,t-d} = \frac{e^{\text{SCORE}_{i,t-d}}}{\sum_{d=22}^{252} e^{\text{SCORE}_{i,t-d}}}, \tag{8}$$

which assures that the boundary conditions outlined in Eq. (3) are fulfilled. A larger weight corresponds to the model assigning a greater importance to the return realized on that day for following the trend in the following month. The model is estimated by minimizing the squared distance between realized returns $r_{i,t+1m}$ and expected returns shown in Eq. (2).[3] We use the Adam optimizer (Kingma and Ba, 2014), a variant of stochastic gradient descent that converges faster and leads to better optima.

To obtain an investable CMM-based portfolio, we follow the standard momentum literature (Jegadeesh and Titman, 1993; Daniel and Moskowitz, 2016) and sort firms available at the end of month t into deciles based on $\mathbb{E}_t^{\text{CMM}}[r_{i,t+1m}]$. Within each portfolio, we value-weight the included stocks. The CMM-based momentum portfolio is then the high-minus-low dollar-neutral portfolio, which we can directly compare to the standard momentum portfolio, with the sole difference being in the way past returns are aggregated to form expected return signals.

3. Investment Performance

Here, we describe the data sources used to conduct our empirical analyses, and investigate the changes in the profitability of momentum-based strategies that arise from our proposed change in how momentum signals are generated.

Data. We obtain daily return data from CRSP and firm characteristics from Jensen et al. (2023) between 1973 and 2022. We consider the 153 firm characteristics studied by Jensen et al. (2023), which are collected in $\boldsymbol{z}_{i,t}$. We follow Gu et al. (2020) and iteratively expand the training sample: for the first iteration, we use data from 1973 through 1982 to estimate the model and validate its fit. The validation period comprises the last 30% of months during that period. We then use the fitted model in an out-of-sample test for

[3] To facilitate the estimation of the machine learning model, we cross-sectionally normalize the target returns in each month. Importantly, this maintains the idea of differentiating winners from losers in a cross-sectional momentum strategy.

Table 1: Performance of Characteristic-Managed Momentum

The table shows performance metrics for both the Characteristic-Managed Momentum (CMM) and standard momentum (SM) strategy. We show both, the full sample metrics spanning 1983 through 2022, as well as results for the pre- and post-2003 sample period. MDD denotes the maximum drawdown of the portfolio's value and CE($\gamma = 5$) denotes the certainty equivalent of a CRRA investor with a risk-aversion level of 5.

	CMM			SM		
	All	Pre 2003	Post 2003	All	Pre 2003	Post 2003
$\overline{r}$	0.255	0.357	0.157	0.226	0.359	0.100
$\sigma(r)$	0.162	0.177	0.141	0.325	0.324	0.322
SR	1.574	2.022	1.116	0.695	1.107	0.310
MDD	−0.276	−0.276	−0.269	−0.832	−0.476	−0.832
Skewness	0.403	0.423	0.053	−1.158	−0.954	−1.414
Kurtosis	3.144	3.475	1.547	4.831	5.203	4.655
CE($\gamma = 5$)	0.191	0.282	0.108	−0.191	−0.045	−0.322

all months of 1983 and 1984. Next, we expand the training and validation data set to encompass data from 1973 – 1984 to test the model's out-of-sample properties in 1985 through 1986. We iterate forward until all years are exhausted. This allows us to study the out-of-sample properties of CMM for the four decades between 1983 and 2022.

Profitability. Table 1 provides a comparison of characteristic-managed momentum (CMM) and standard momentum (SM). The sole difference between the two strategies is how the momentum signal – the aggregation of past daily returns – is constructed. The portfolio construction is the same, in that we show the high-minus-low return spread of value-weighted decile portfolios. CMM generates an average annualized return of 25.5%, compared to 22.6% for SM over the testing sample. The standard deviation of CMM is significantly lower at 16.2% per year vs. 32.5% for SM, which culminates in an annualized Sharpe ratio of CMM more than twice as large as that of SM (1.57 vs. 0.70). The returns of CMM are also considerably more stable over time, with a maximum drawdown of −27.6%, compared to −83.2% for momentum during the financial crisis of 2008/2009. Flexibly weighting past returns circumvents the issue of momentum crashes (Daniel and Moskowitz, 2016). A positive skewness of CMM corroborates this finding. For a moderate risk aversion of $\gamma = 5$, the certainty equivalent (CE) return of a CRRA-utility investor amounts to 19.1% per year. In contrast, the same investor would have foregone an investment in standard momentum with a negative CE, as a result of the strategy's significant volatility and drawdown.

Table 2: Independent Double Sort

The table shows time-series average returns of portfolios independently sorted into quintiles by the standard momentum signal (SM) and by CMM's weighted past return signal. Within each portfolio we value-weight stocks with their market capitalization. We also show high-minus-low returns. *** (**, *) denotes that the return is statistically significant at the 1% (5%, 10%) level using HAC robust standard errors with 12 lags.

	CMM					
SM	LOW	2	3	4	HIGH	HL
LOW	-0.91^{*}	0.29	0.24	0.79	1.41^{***}	2.32^{***}
2	-0.26	0.49	0.42	0.87^{***}	1.24^{***}	1.51^{***}
3	-0.29	0.22	0.63^{**}	0.89^{***}	1.30^{***}	1.58^{***}
4	-0.10	0.57^{*}	0.68^{***}	0.91^{***}	1.14^{***}	1.24^{***}
HIGH	0.19	0.88^{**}	0.87^{***}	1.02^{***}	1.32^{***}	1.14^{***}
HL	1.10^{*}	0.59	0.62	0.23	-0.09	-1.19^{***}

Green et al. (2017) show that many characteristic-sorted investment strategies fail to perform after 2003. We therefore perform a sample split to investigate if CMM – unlike other characteristic-based investment strategies – continues to work in recent decades. Consistent with a performance attenuation, we find that SM generates much smaller returns since 2003 of just 10% per year, at an annualized Sharpe ratio of only 0.31, underperforming a buy and hold investment in the market for the same period. In contrast, CMM continues to deliver: average annualized returns amount to 15.7% with a Sharpe ratio of 1.12, tripling that of standard momentum and outperforming the market. While CMM's performance also attenuates after 2003, it continues to deliver profitable investment advice, highlighting that momentum-type trading and trend-following in general is still profitable today, but requires a flexible weighting of past returns to guide the decision to invest.

CMM Subsumes Standard Momentum. Given the drastically improved performance of CMM compared to standard momentum, we investigate in Table 2, whether CMM subsumes the profitability of standard momentum. If it does, this highlights that changing how past daily returns are weighted not only outperforms an imposed equal importance of past returns, but that the deviation in weights from standard momentum is a strictly better way of composing momentum-based signals. Each month, we independently sort stocks into quintiles based on the standard momentum signal and CMM's signal, and record the average realized returns of the 25 portfolios, as well as the corresponding high-minus-low portfolios. We continue to value-weight stocks with their

market capitalization.

We find that the CMM signal fully subsumes the profitability of the standard momentum signal. Within each of the five CMM portfolios, the resulting high-minus-low portfolio sorted by the standard momentum signal fails to generate average returns that are significant at the 5%-level. In contrast, within each of the five standard momentum portfolios, we find an economically large and statistically significant CMM-based return spread, which ranges from 1.14% per month for momentum winners (SM-HIGH) to 2.32% for momentum losers (SM-LOW). Not only do these results show the unequivocal improvement of allowing the weight placed on individual past daily returns to differ, but they also highlight that CMM particularly improves upon standard momentum by identifying those previous losers that are expected to rebound and outperform the remaining stocks. This culminates in the lowered maximum drawdown evident in both Table 1 and Figure 5, which we analyze in more detail in Section 4.2.

Which Days Are Important for Momentum? To understand how and why CMM outperforms standard momentum, it is instructive to investigate the output of the weighting function and to see how weights w differs from an equal importance assigned by standard momentum. We first show that it is sparse on average, suggesting that some days are significantly more important than others. We then also show that earnings announcements, and days with particularly large returns are most important in the formation of CMM's signal.

Figure 2 plots the cumulative average weight of CMM against the equal weights proposed by standard momentum. The figure can be understood as an empirical cumulative density function of weights w put on past returns. On average, just two of the 231 days in the formation period, already receive a weight of 30%, 35 days already account for more than half of the weights on average, highlighting a large degree of sparsity in the importance assigned by CMM to individual return days. We also document a large dispersion around this mean: the shaded area shows the variation within the 25th and 75th percentile of this cumulative weighting function across all firm×month observations in our out-of-sample period. The weighting of several past daily returns is significantly more sparse than the average, while several others revert back to standard momentum's equal importance.

As a natural follow-up, we investigate the average importance of certain event and news days. In each month, we perform cross-sectional regressions of the weights placed

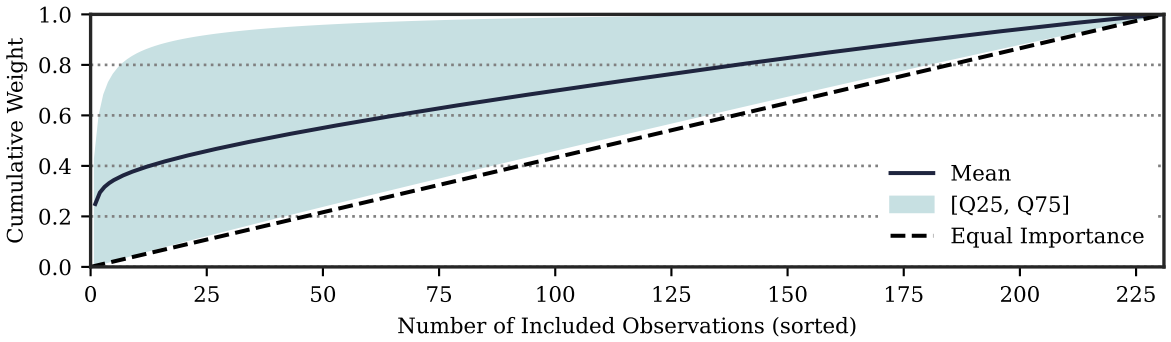


Fig. 2. Sparsity in Weights

The figure shows cumulative weights sorted in descending order of the average weighting function ω. The shaded area highlights the variation within the 25% and 75% quantile. On average, 1% of the past observations account for 30% of the weights; 15% of observations already account for about 50% of the weights on average.

on daily returns in the formation period on different news-day indicators. We consider earnings and FOMC announcement days, the announcements of GDP and CPI figures, and days with large jumps in the stock market, which Baker et al. (2021) identify. We also investigate if the returns of the more recent months in the formation period (months $t-6m$ to $t-2m$) or months of the latter half ($t-12m$ to $t-7m$) are more important for the generation of CMM signals, following the insight of Novy-Marx (2012) that momentum signals based on the intermediate horizon are more profitable. Finally, we investigate if days in the formation period with particularly large returns are more important, focusing on the 5% of days with the largest absolute returns. The regressions take the following form:

$$w_{i,t-d} = \alpha + \beta \cdot \mathbb{1}_{d=x} + \epsilon_{i,t-d}, \tag{9}$$

where $\mathbb{1}$ is the indicator function and x represents one of the aforementioned news and event days.

The results in Table 3 produce several interesting findings. We first find that earnings announcement days are of particular importance. On average, they are weighted 115% higher than the average day, highlighting that the return response to earnings announcement carries important information for future returns.[4] Weights placed on FOMC announcement days are 15% larger than the average, whereas both weights on GDP and CPI announcement days are not distinguishable from the average weight. When we combine the macroeconomic and earnings announcement days in Panel B, however, we

[4]The coefficients in Table 3 show the relative increase in the average weight.

Table 3: Two-Stage Regression

The table shows results of a two-stage regression analysis. For each date we perform cross-sectional regression of daily weights on different news-days indicators (cf. Eq. (9)). We then run time-series regression to obtain estimates of the importance of these news days relative to assigning equal weights as in standard momentum. We use HAC-robust standard errors with 12 lags to calculate t-values. News days we consider are earnings announcement days (EA) and days with large jumps in the stock market (Baker) as defined by Baker et al. (2021). We also add to the discussion on whether the most recent-past returns (r_6_1) are more important relative to the intermediate horizon (Novy-Marx, 2012). Finally, we consider the weights on past large returns , that is, the 5% of days with the largest absolute return ($|r| > Q95$) within each time-series. To account for possible pre- and post-announcement drifts or after-hour announcements we use the average weight within a 3 day window from $t-1$ to $t+1$ days around the announcement date for EA, CPI and GDP.

	EA	Baker	FOMC	GDP	CPI	r_6_1	$\|r\| > Q95$
				Panel A: Univariate Regressions			
Coeff.	1.15	1.27	0.15	−0.00	−0.01	0.00	5.42
$t-$value	(13.81)	(5.14)	(2.43)	(−0.23)	(−1.49)	(0.02)	(25.18)
				Panel B: Multivariate Regression			
Coeff.	1.14		0.10	−0.03	−0.03		
$t-$value	(13.24)		(1.56)	(−2.19)	(−3.37)		

find that the coefficient of the earnings announcements does not change, whereas the overweighting of FOMC announcements now turns insignificant, and both GDP and CPI announcement days are significantly underweighted. As macroeconomic and firm-specific news enters the market on the same day, the importance placed on macroeconomic releases drops, with an unchanged and large importance assigned to the reaction to earnings news. This echoes some of the findings by Hirshleifer and Sheng (2022) who show that the sensitivity of stock returns to earnings announcements is elevated on days with an accompanying macroeconomic news announcement.

In the remaining regression specifications we show that both returns in the first five months of the formation period, and the remaining six are equally important on average. Days on which the overall stock market jumps are overweighted by 127% on average, and days with a particularly large return of the stock in question are significantly more informative of future returns, being weighted 542% larger than an equal weighting. This is consistent with market microstructure theory, which postulates that price innovations are the consequence of new information (e.g. Kyle, 1985), such that larger price innovations carry more information. CMM is able to exploit the information released on these days.

In Appendix A we show that constructing equally-weighted momentum signals from the

aforementioned news and event days is insufficient to construct a momentum strategy that systematically outperforms standard momentum and CMM, highlighting that a flexible weighting of the information contained in the return reaction to these days is vital to CMM's success.

4. What Creates Momentum?

Our study is chiefly concerned with understanding why momentum profits arise in the stock market. Whereas the observation that momentum profits have been severely muted since 2003 (Green et al., 2017) may question the importance of studying the momentum anomaly, the previous section has shown that momentum-type strategies are alive and well, if the investor acknowledges that past realized returns differ in how much information they convey. The simple change in how past realized returns are weighted to construct momentum signals delivers profitable investment advice also in the most recent decades. We use this insight to differentiate possible explanations for the success of the CMM momentum strategy.

Although momentum in the cross-section of stocks has been studied for more than thirty years, going back to the seminal paper of Jegadeesh and Titman (1993), there is still no consensus about the origins and drivers of momentum profits in that market. Possible explanations can be broadly clustered into two "camps", one relating momentum profits to investors' biases in how they react to news, the second to their risk exposure.

For example, momentum profits may be explained through the lens of an over- or underreaction to news. In the model of Daniel et al. (1998), investors are overconfident in the signals they generate themselves and will disregard conflicting public information. This pushes prices up, leading to momentum. Over time, as public information is gradually impounded into market prices, the stock's value will fall back in line with its fundamental value. Advocating for a story of underreacting to public news, Barberis et al. (1998) model investors as conservative, leading them to underweight new information. The subsequent adjustment of the stock's price to reflect its fundamental value generates a positive autocorrelation in returns. Instead, Kelly et al. (2021) show that past return-based characteristics predict future realized betas, and show that conditional risk exposure can explain momentum profits. Also favoring a risk-based explanation, Daniel and Moskowitz (2016) show that momentum profits are plagued by occasional but very

severe crashes, which take decades to recover.

To assess which of these explanations best aligns with the profitability of CMM in the stock market, we first show by analyzing the weights w assigned by CMM that an underreaction to news released in the formation period is a better explanation than an overreaction to news. We then conduct a variety of additional empirical tests for risk-based and behavioral explanations, which are inspired by the literature investigating the origins of momentum profits.

4.1. Under- or Overreaction to News

Which Past Returns are Overweighted? To see how the trajectory of weights w can informs us about the relative likelihood that realized momentum profits are the result of a systematic over- or a systematic underreaction to the news released during the formation period, assume that an important information event falls on day $t - d*$. If market participants *underreact* to this news and the news itself is positive for the firm, the realized return on $t - d*$ is too small, such that $r^{\text{realized}}_{t-d*} > 0$ and $r^{\text{realized}}_{t-d*} < r^{\text{efficient}}_{t-d*}$. If instead the news is negative for the firm, we have that $r^{\text{realized}}_{t-d*} < 0$ and $r^{\text{realized}}_{t-d*} > r^{\text{efficient}}_{t-d*}$. In both cases, the absolute realized return is closer to zero than what would have been efficient, as judged by the news component released on that day. To capture the resolution of this mispricing during the investment period, CMM should thus assign a greater weight to r_{t-d*}, as market prices are expected to align with fundamental values. As a result, CMM will capture an underreaction to news if the correlation between past absolute returns $|r|$ and the weights w assigned to them is positive. Analogously, to profit from a systematic overreaction of market participants to the news, the correlation between past absolute returns and weights will be negative. This insight provides us with a simple to use tool to understand whether trend-following profits are primarily driven by market participants under- or overreacting to news. Importantly we can quantify the perceived share of under- vs. overreaction in the cross-section of stocks.

For all firm×month observations in our sample, we compute the corresponding correlation coefficient between past absolute daily returns and the weights placed on them by the CMM model. We also compute the coefficient's statistical significance. CMM extrapolates from an underreaction (overreaction) to news during the formation period, whenever the coefficient is positive (negative) and statistically significant at the 1%-level.

Figure 3 shows the average fraction of the cross-section of stocks per year, for which we

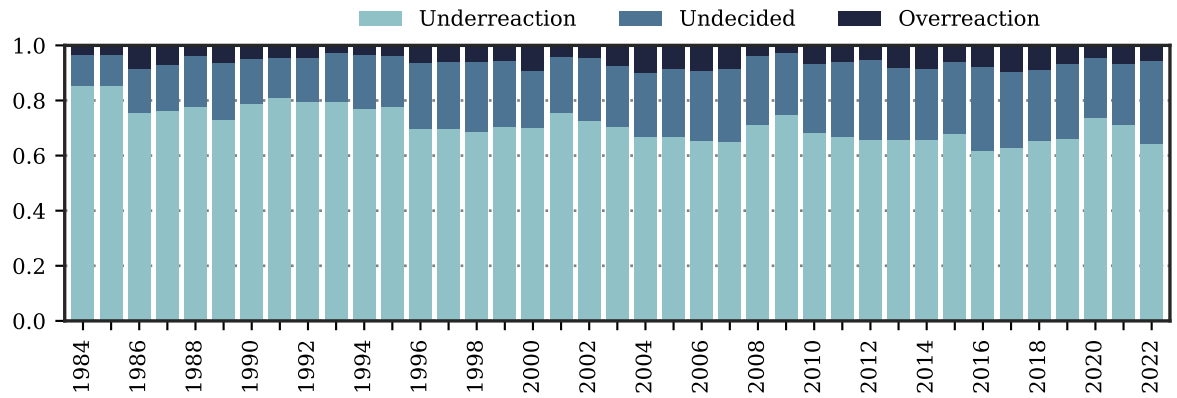


Fig. 3. Under- vs. Overreaction

The figure shows the fraction of firm×month observations, for which the model implied correlation coefficient between the weights $w_{i,t-d}$ and absolute returns $|r_{i,t-d}|$ across all lookback days d is significantly positive or negative for different vintage years. A significantly positive correlation coefficient suggests that the model positively extrapolates past absolute returns, in line with an initial underreaction to the news component in past returns, while a negative correlation suggests the opposite. The verdict is "undecided" whenever the correlation coefficient is not statistically different from zero at the 1%-level.

identify a systematic underreaction to news in light blue and a systematic overreaction to news in black. For the remaining firm×month observations, the correlation between absolute past daily returns and weights is statistically indistinguishable from zero. The results are clear: throughout our out-of-sample period of 40 years, we find a clear dominance of CMM exploiting an underreaction to news. For between 63% and 87% of firms does the model generate a positive correlation between past absolute daily returns and weights, consistent with this explanation. An overreaction, in contrast, is rare, being found for fewer than 10% of the stocks in the average month. For the remainder of the cross-section, we find that the correlation between absolute returns and weights is insignificant, such that the extracted CMM signals on average collapse to those of standard momentum.

To further ascertain that the dominant explanation for momentum-based profits in the market stems from an underreaction to news, we separately investigate the profitability of the CMM strategy for firms within each of the three buckets depicted in Figure 3, capturing an under- (overreaction) to news, or an undecided verdict. The results are shown in Table 4. Consistent with the breakdown in Figure 3 showing that CMM predominantly extracts a trend from an underreaction to news released in the formation period, its trading signal among these stocks is most valuable: the value-weighted high-minus-low quintile spread amounts to 1.73% per month, which is highly statistically significant. At

Table 4: CMM's Performance and Correlation

The table shows time-series average portfolio returns. Stocks are first sorted by the correlation of their past absolute returns $|r|$ and weight w allocated by our model. Within each portfolio stocks are then again sorted into quintiles by CMM's signal. We continue to value-weight stocks with their market capitalization within each portfolio. *** (**, *) denotes that the return is statistically significant at the 1% (5%, 10%) level using HAC robust standard errors with 12 lags.

	CMM					
	LOW	2	3	4	HIGH	HL
Overreaction	0.47	0.63	0.48	0.77**	1.09***	0.62
Undecided	0.29	0.53	0.36	0.85***	0.69*	0.41
Underreaction	−0.42	0.51	0.90***	0.95***	1.31***	1.73***

the same time, we find that that CMM has a harder time exploiting an overreaction. The high-minus-low return spread is not statistically significant, but nonetheless economically meaningful at 0.62% per month. Of course, the circumstance that the return spread is not statistically distinguishable from zero could also reside in the fact that CMM rarely identifies stocks for which exploiting an overreaciton is optimal. Fewer than 10% of firm×month observations are identified as such. Finally, we find corroborating evidence that assigning an equal importance to the returns realized in the formation period is supoptimal. Whenever CMM's weighting of past returns collapses to that of standard momentum, evident in a correlation between absolute returns and weights that is indistinguishable from zero, the resulting high-minus-low return spread is lowest at 0.41% per month, which is not statistically significant. Not only do these results favor a story of underreacting to news as the source of momentum profits, but they also highlight the economic gains achieved when extracting an optimal return weighting function from the data.

Long-term Performance. The results in Figure 4 corroborate the intuition that a systematic underreaction to news is a likely explanation for momentum profits in the stock market. The figure shows the returns of the value-weighted high-minus-low decile portfolio based on CMM, when the expected return signal is generated at the end of month t and the portfolio return is evaluated in month $t + \tau$ with $\tau \in [1, 24]$ months. We find a gradual decline in the average realized returns, as we increase τ. A systematic overreaction to news released in the formation period should lead to an initially positive performance of CMM but a correction of this return pattern in the long-run, as market

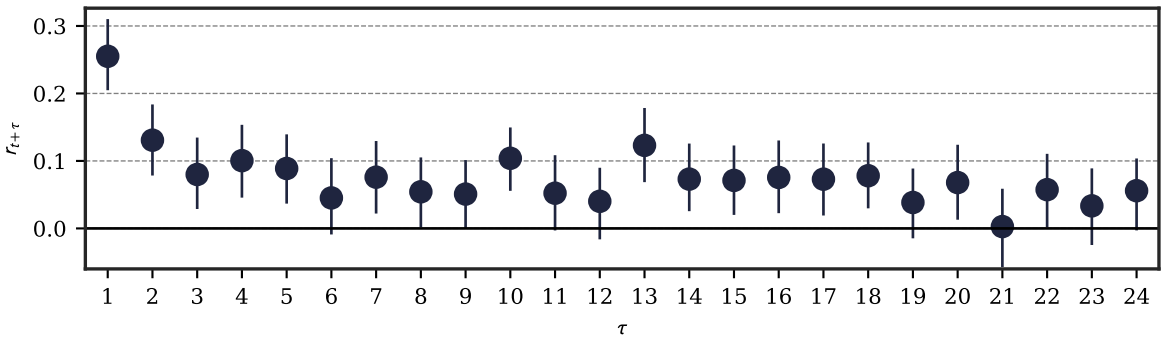


Fig. 4. Average Returns of CMM Formed on Lagged Signals

The figure shows time-average returns of the CMM portfolio formed on τ-periods lagged signals. That is, we keep the signal generated in time t fixed and record returns of the CMM portfolio in month $t+\tau$, with $\tau \in [1, 24]$.

prices align with fundamentals.[5]

4.2. Risk-based Explanations

Time-varying Risk Compensation. While our results up to this point favor a systematic underreaction to news as the driving force behind momentum-based profits in the stock market, CMM's signal may also capture time-variation in the compensation for risks. To this end, Kelly et al. (2021) show that the predictive power of momentum is subsumed by a measure of the conditional compensation for risk. Specifically, they estimate a conditional latent factor model of the form:

$$r_{i,t+1m} = \beta'_{i,t} f_{t+1m} + \varepsilon_{i,t+1m}, \qquad \mathbb{E}_t[r_{i,t+1m}] = \beta'_{i,t}\lambda_t, \tag{10}$$

and show that $\beta'_{i,t}\lambda_t$ is a better predictor of future stock returns than the momentum signal and subsumes its predictability. Their rationalization of this finding is that past returns, used to form momentum signals, are a proxy for the time-variation in βs, which in turn suggests that the profitability of momentum strategies is the compensation for bearing systematic risk.

We follow Kelly et al. (2021) and approximate the model in Eq. (10) using the instrumented principal component analysis (IPCA) put forth by Kelly et al. (2019) with five

[5]In untabulated results, we confirm that the returns converge to and stay at zero for $\tau \in [25, 60]$ months.

latent factors f:

$$r_{i,t+1m} = \beta'_{i,t} f_{t+1m} + \varepsilon_{i,t+1m}, \qquad \beta_{i,t} = z'_{i,t}\Gamma_\beta + \nu_{i,t}. \tag{11}$$

Here, βs are a linear function of observable firm characteristics z. We use the same set of 153 firm characteristics as in the estimation of the CMM model. We estimate the IPCA model in a rolling out-of-sample fashion and estimate the price of risk as the average of factor realizations in the estimation period, i.e., $\widehat{\lambda}_t = \frac{1}{t}\sum_{\tau=0}^{t} f_\tau$.

In panel regressions, we then investigate if the predictive power over future monthly excess stock returns of the CMM-based signal is subsumed by the time-varying compensation for risk, $\beta_{i,t}\lambda_t$:

$$r_{i,t+1m} = c_0 + c_1 \cdot \mathbb{E}_t^{\text{CMM}}[r_{i,t+1m}] + c_2 \cdot (\beta_{i,t}\lambda_t) + \epsilon_{i,t+1} \tag{12}$$

The results are provided in Table 5. As in Kelly et al. (2021), we separately show the results for the raw momentum signals in the columns to the left, and the cross-sectional percentage ranks of the signals in the columns to the right.

For both the raw signals and their cross-sectional ranks, we find in specification (1) that the CMM-based momentum signal predicts future stock returns well. Stocks with a higher CMM signal earn significantly larger returns over the next month. We next confirm in specification (2) the result of Kelly et al. (2021) that a description of the time-varying compensation for risks in the stock market also predicts future stock returns well: both the coefficient for the raw $\beta'\lambda$ signal and its cross-sectional ranks are large and

Table 5: CMM and Time-Varying Risk

The table shows univariate and bivariate panel regressions results. We report coefficient estimates and corresponding t-statistics using standard errors clustered by date. Coefficients for regressions using ranks are multiplied by 100 for readability.

	Raw			Ranks		
	(1)	(2)	(3)	(1)	(2)	(3)
Const.	0.01	0.00	0.00	0.01	0.01	0.01
	(2.61)	(0.25)	(1.12)	(2.57)	(2.57)	(2.57)
CMM	0.09		0.06	3.72		2.27
	(14.33)		(9.68)	(15.63)		(8.00)
$\beta'\lambda$		0.80	0.50		3.69	2.18
		(16.25)	(11.70)		(16.67)	(8.50)

statistically highly significant.[6] Finally, in specification (3), we combine both signals in a single regression. We find that while both coefficients decrease by about a third, compared to in the univariate regressions, they retain their economic and statistical significance. Whereas the predictive power of standard momentum is well explained by it capturing time-variation in the compensation for risks in the stock market, the predictability of CMM's signal is only modestly impacted.[7] The absence of a clear relationship between time-varying risk compensation and CMM-based momentum profits in turn gives further credence to our empirical results in favor of a systematic underreaction to news giving rise to momentum in the stock market.

Momentum Crashes. Daniel and Moskowitz (2016) propose an explanation for momentum profits in their inherent risk of so-called "momentum crashes" – brief investing periods that wipe out decades of previous profits. Our results in Table 1 and the visualization of the cumulative dollar value of investing in CMM in Figure 5 suggest that CMM is not subject to the same severity of crashes as standard momentum is. In fact, the maximum drawdown for the full sample period is -27.6%, which compares favorably to the -83.2% lost by standard momentum. In turn, our results suggest that while crashes are a potential explanation for standard momentum, they do not serve as a comprehensive explanation for all types of momentum-based profits.

[6]Note that the size of the coefficients for raw signals are not comparable as CMM targets cross-sectionally standardized returns as explained in Section 2.

[7]We replicate the results of Kelly et al. (2021) that $\beta'\lambda$ subsumes the predictive power of the standard momentum signal in Appendix B.

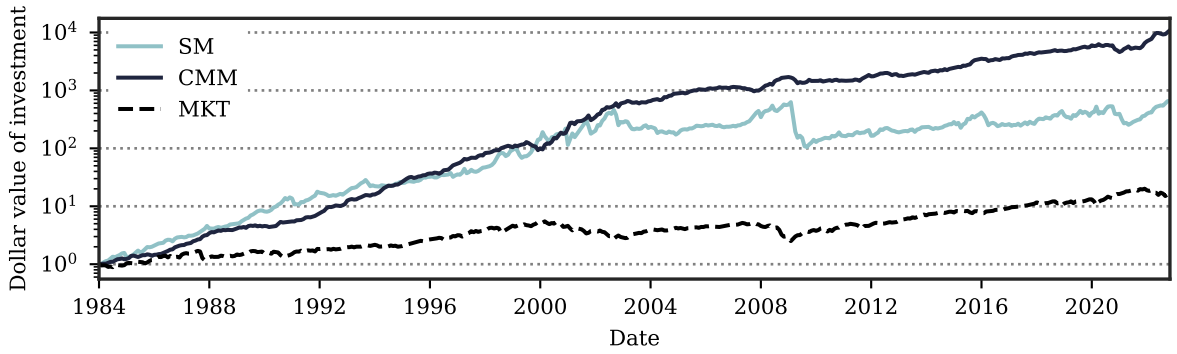


Fig. 5. Performance of Characteristic-Managed Momentum

The figure shows the cumulative returns to the Characteristic-Managed Momentum strategy (CMM) and to standard momentum (SM) for the testing sample from 1983 through 2022 on a log scale. For comparison, we also add the cumulative returns of the market portfolio (MKT) to the graph.

The results of the double sort in Table 2 furthermore show that the resilience of CMM to momentum crashes lies in the identification of stocks that are classified as "losers" by standard momentum, but that significantly outperform over the coming months. Among standard momentum losers, the CMM signal is most profitable, and generates a highly significant value-weighted high-minus-low decile spread of 2.3% a month.

5. Robustness

In this section, we rule out potential alternative explanations for CMM's profitability. We find that the "frog-in-the-pan" hypothesis of Da et al. (2014) cannot explain its success, suggesting that information discreteness alone is an insufficient explanation of momentum-type profits. We also show that leading stock market factor models are unable to explain the returns produced by CMM, and that alternative augmentations of the standard momentum strategy, which have been proposed by the literature, fail to match CMM's performance.

Information Discreteness. Da et al. (2014) show that momentum profits are amplified substantially, whenever the information embedded in the daily returns of the formation period is impounded gradually, instead of in sudden jumps. The authors attribute this to limited attention by investors placed on gradual changes, whereas jumps are salient. In Appendix C we first confirm their finding that standard momentum is significantly elevated when the information in the formation period is impounded continuously. Average momentum profits in this group of stocks amount to 1.77% per month, which pales the insignificant returns of 0.42% among stocks with a discrete impounding of information. In contrast, however, we find no evidence that the discreteness with which information is impounded throughout the formation period impacts the profitability of CMM. Instead, the high-minus-low return spread is economically large and statistically significant regardless of whether the information is released continuously or in discrete chunks.

Factor Exposure. We also find no evidence that CMM's exposure to common factors explains its performance. Appendix D shows that the strategy produces economically large and statistically significant αs for the Fama and French (2015) five factor model augmented by standard momentum, the q-factor model by Hou, Xue, and Zhang (2015),

and the behavioral factor model of Daniel, Hirshleifer, and Sun (2020). αs range between 1.6% for the q-factor model to 2% for the behavioral factor model.

Alternative Augmentations of Momentum. As a final robustness check, we compare CMM's performance improvement over standard momentum to that of other strategies based on different signal definitions proposed in the literature. We consider momentum strategies based on intermediate horizon past performance (MOM_12_7) and recent past performance (MOM_6_1), following Novy-Marx (2012) as well as the price-to-high momentum measure by George and Hwang (2004). None of these signals achieve a performance comparable to that of CMM.

6. Conclusion

While the traditional momentum strategy assumes that returns realized in the formation period are of equal importance, we show that allowing for a flexible aggregation of these returns generates significantly more profitable momentum signals. The resulting characteristic-managed momentum strategy subsumes the performance of standard momentum, continues to perform well after 2003, and is resistant to momentum crashes, as it correctly identifies those previous losers that outperform in the future.

We find that the weights produced by our simple machine learning framework are sparse: just 35 of the 231 daily returns observed in the formation period already receive more than 50% of the total weight on average. Earnings announcements, days of market-wide jumps, and days with particularly large stock-specific returns are most informative.

Finally, we show that the correlation between the weights and absolute daily returns informs us about whether an over- or an underreaction to the information released throughout the formation period is a more likely candidate for explaining the success of momentum strategies in the stock market. Our results strongly favor an underreaction to news as the dominant driver of CMM's success: the strategy exploits an underreaction to news for between 60-90% of stocks and only within this subsample do we find a statistically significant and economically meaningful value-weighted high-minus-low return spread. We perform several empirical tests to rule out alternative behavioral and risk-based explanations. Importantly, we show that CMM's profits are not explained by the variation in the compensation for risks in the stock market.

References

Asness, C. S., Moskowitz, T. J., Pedersen, L. H., 2013. Value and momentum everywhere. The Journal of Finance 68, 929–985.

Baker, S. R., Bloom, N., Davis, S. J., Sammon, M. C., 2021. What triggers stock market jumps? Working Paper, National Bureau of Economic Research.

Bali, T. G., Beckmeyer, H., Moerke, M., Weigert, F., 2023. Option return predictability with machine learning and big data. Review of Financial Studies 36, 3548–3602.

Bali, T. G., Goyal, A., Huang, D., Jiang, F., Wen, Q., 2020. Predicting corporate bond returns: Merton meets machine learning. Working Paper, Available at SSRN.

Barardehi, Y. H., Bogousslavsky, V., Muravyev, D., 2022. What drives momentum and reversal? Evidence from day and night signals. Working Paper, Available at SSRN 4069509.

Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. Journal of Financial Economics 49, 307–343.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Chen, L., Pelger, M., Zhu, J., 2023. Deep learning in asset pricing. Management Science pp. 1–37.

Da, Z., Gurun, U. G., Warachka, M., 2014. Frog in the pan: Continuous information and momentum. The Review of Financial Studies 27, 2171–2218.

Daniel, K., Hirshleifer, D., Subrahmanyam, A., 1998. Investor psychology and security market under-and overreactions. The Journal of Finance 53, 1839–1885.

Daniel, K., Hirshleifer, D., Sun, L., 2020. Short-and long-horizon behavioral factors. The Review of Financial Studies 33, 1673–1736.

Daniel, K., Moskowitz, T. J., 2016. Momentum crashes. Journal of Financial Economics 122, 221–247.

Ehsani, S., Linnainmaa, J. T., 2022. Factor momentum and the momentum factor. The Journal of Finance 77, 1877–1919.

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

Ernst, R., Gilbert, T., Hrdlicka, C. M., 2019. More than 100% of the equity premium: How much is really earned on macroeconomic announcement days? Working Paper, Available at SSRN 3469703.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

George, T. J., Hwang, C.-Y., 2004. The 52-week high and momentum investing. The Journal of Finance 59, 2145–2176.

Green, J., Hand, J. R., Zhang, X. F., 2017. The characteristics that provide independent information about average us monthly stock returns. The Review of Financial Studies 30, 4389–4436.

Gu, S., Kelly, B., Xiu, D., 2020. Empirical asset pricing via machine learning. The Review of Financial Studies 33, 2223–2273.

Heston, S. L., Jones, C. S., Khorram, M., Li, S., Mo, H., 2023. Option momentum. The Journal of Finance, 78, 3141–3192.

Hirshleifer, D., Sheng, J., 2022. Macro news and micro news: complements or substitutes? Journal of Financial Economics 145, 1006–1024.

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

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

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

Jensen, T. I., Kelly, B. T., Pedersen, L. H., 2023. Is there a replication crisis in finance? The Journal of Finance 78, 2465–2518.

Kelly, B. T., Moskowitz, T. J., Pruitt, S., 2021. Understanding momentum and reversal. Journal of financial economics 140, 726–743.

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

Kingma, D. P., Ba, J., 2014. Adam: A method for stochastic optimization. Working Paper, arXiv preprint arXiv:1412.6980.

Knox, B., Londono, J. M., Samadi, M., Vissing-Jorgensen, A., 2024. Equity Premium Events. Working Paper, Available at SSRN.

Kyle, A. S., 1985. Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society pp. 1315–1335.

Li, S. Z., Yuan, P., Zhou, G., 2023. Risk-based momentum of corporate bonds. Working Paper, Available at SSRN 4374766.

Liang, L., 2003. Post-earnings announcement drift and market participants' information processing biases. Review of Accounting Studies 8, 321–345.

Lou, D., Polk, C., Skouras, S., 2019. A tug of war: Overnight versus intraday expected returns. Journal of Financial Economics 134, 192–213.

Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Currency momentum strategies. Journal of Financial Economics 106, 660–684.

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

Novy-Marx, R., 2012. Is momentum really momentum? Journal of Financial Economics 103, 429–453.

Savor, P., Wilson, M., 2013. How much do investors care about macroeconomic risk? evidence from scheduled economic announcements. Journal of Financial and Quantitative Analysis 48, 343–375.

Savor, P., Wilson, M., 2016. Earnings Announcements and Systematic Risk. The Journal of Finance 71, 83–138.

van Binsbergen, J. H., Han, X., Lopez-Lira, A., 2023. Man vs. machine learning: The term structure of earnings expectations and conditional biases. The Review of Financial Studies 36, 2361–2396.

Appendix

A. Momentum Signals From News and Event Days

The table shows performance metrics for alternative momentum trading strategies based on different signals. Signals are based on the average return for different news days within the lookback period of $t-12$ and $t-1$ months. For earnings announcements (EA), GDP, CPI we consider a three day window around the news release date to account for announcement drifts and after-hour news releases. BAKER refers to large jumps in the stock market as defined by Baker et al. (2021) and $Q5$ considers the 5% days within the lookback-window with the highest absolute daily returns. Ensemble denotes an ensemble average of all signals except CMM. If there happen to be no news dates available for any stock, we set the signal to zero. In the table, MDD denotes the maximum drawdown of the portfolio's value and CE($\gamma = 5$) denotes the certainty equivalent of a CRRA investor with a risk-aversion level of 5. All portfolios are value-weighted.

	CMM	EA	FOMC	GDP	CPI	BAKER	Q5	Ensemble
$\overline{r}$	0.255	0.045	0.112	0.101	0.125	−0.052	0.007	0.076
$\sigma(r)$	0.162	0.104	0.173	0.199	0.198	0.192	0.254	0.210
SR	1.579	0.433	0.647	0.507	0.631	−0.273	0.029	0.362
MDD	−0.276	−0.479	−0.528	−0.612	−0.461	−0.946	−0.903	−0.731
Skewness	0.401	0.114	0.595	−0.739	−0.786	−1.392	1.177	−0.036
Kurtosis	3.152	2.666	6.092	3.508	4.040	10.169	11.346	4.688
CE($\gamma = 5$)	0.192	0.018	0.038	−0.012	0.012	−0.179	−0.154	−0.042

B. Momentum and Time-Varying Risk Compensation

The table shows univariate and bivariate panel regressions results (cf. Eq. (12)). We report coefficient estimates and corresponding t-statistics using standard errors clustered by date. Coefficients for regressions using ranks are multiplied by 100 for readability.

	Raw			Ranks		
	(1)	(2)	(3)	(1)	(2)	(3)
Const.	0.01	0.00	0.00	0.01	0.01	0.01
	(2.48)	(0.25)	(0.25)	(2.57)	(2.57)	(2.57)
SM	0.00		−0.00	1.30		0.51
	(0.61)		(−0.09)	(3.35)		(1.20)
$\beta'\lambda$		0.80	0.81		3.69	3.58
		(16.25)	(15.58)		(16.67)	(13.20)

C. Information Discreteness and Momentum Profits

The table shows time-series average returns of portfolios sorted first into quintiles by our CMM return forecast (left) or standard momentum (right) and subsequently by the information discreteness measure ID of Da et al. (2014). Stocks within each portfolio are weighted by their market capitalization. *** (**, *) denotes that the average portfolio return is statistically significant at the 1% (5%, 10%) level calculated using HAC-robust standard errors with 12 lags.

	CMM			SM		
Information	LOW	HIGH	HL	LOW	HIGH	HL
Discrete	−0.27	1.08***	1.36***	0.18	0.60*	0.42
2	−0.43	1.10***	1.53***	−0.12	0.87***	0.99**
3	−0.26	1.35***	1.61***	−0.43	0.95***	1.38***
4	−0.19	1.21***	1.39***	−0.05	0.88***	0.92*
Continuous	−0.32	1.30***	1.62***	−0.70	1.08***	1.77***
Δ	−0.05	0.22	0.27	−0.87***	0.48**	1.35***

D. Factor Exposure

The table shows the results of regressing the returns of CMM on the five factors Fama and French (2015) plus momentum in Panel A, the five factors of the q-factor model of Hou et al. (2015) in Panel B, and the three behavioral factors of Daniel et al. (2020) in Panel C. The factors are obtained from Kenneth French's website, from Kent Daniel's website, and from global-q.org. t-values are given in parentheses and based on Newey and West (1987) standard errors with 12 lags.

	Panel A: FF5 + Momentum							
	α	MKT	SMB	HML	RMW	CMA	MOM	R^2
CMM	0.018	−0.057	−0.051	0.263	0.494	0.198	0.136	0.208
	(7.11)	(−1.17)	(−0.44)	(2.53)	(3.94)	(1.20)	(1.52)	

	Panel B: HXZ q-Factors						
	α	MKT^q	Size^q	IA^q	ROE^q	EG^q	R^2
CMM	0.016	−0.025	0.039	0.525	0.362	0.349	0.186
	(4.83)	(−0.46)	(0.39)	(3.30)	(2.30)	(2.36)	

	Panel C: DHS Behavioral Factors				
	α	MKT	PEAD	FIN	R^2
CMM	0.020	−0.045	−0.072	0.419	0.157
	(6.67)	(−0.86)	(−0.54)	(5.95)	

E. Other Augmentations of Momentum Signals

The table shows full-sample performance metrics for momentum strategies based on different signals. We show results for the Characteristic-Managed Momentum (CMM) and for the standard momentum (SM) strategy. We also consider momentum strategies based on intermediate horizon past performance (MOM_12_7) and recent past performance (MOM_6_1), following Novy-Marx (2012), as well as the price-to-high momentum measure of George and Hwang (2004) (GH04). In terms of performance metrics, MDD denotes the maximum drawdown of the portfolio's value and CE($\gamma = 5$) denotes the certainty equivalent of a CRRA investor with a risk-aversion level of 5.

	CMM	SM	MOM_12_7	MOM_6_1	GH04
$\overline{r}$	0.255	0.226	0.164	0.182	0.085
$\sigma(r)$	0.162	0.325	0.232	0.291	0.237
SR	1.574	0.695	0.705	0.627	0.360
MDD	−0.276	−0.832	−0.600	−0.801	−0.733
Skewness	0.403	−1.158	−0.669	−1.005	−2.177
Kurtosis	3.144	4.831	2.967	6.784	15.038
CE($\gamma = 5$)	0.191	−0.191	0.007	−0.150	−0.260