Momentum has its Moments*

Pedro Barroso† and Pedro Santa-Clara‡

First version: April 2012
This version: April 2013

Abstract

Compared to the market, value or size factors, momentum has offered investors the highest Sharpe ratio. However, momentum has also had the worst crashes, making the strategy unappealing to investors with reasonable risk aversion. We find that the risk of momentum is highly variable over time and quite predictable. The major source of predictability is not time-varying market risk but rather momentum-specific risk. Managing this risk virtually eliminates crashes and nearly doubles the Sharpe ratio of the momentum strategy. Risk management works because high risk forecasts both high risk and low returns. As a result, risk-managed momentum is a much greater puzzle than the original version.

---

*We thank Abhay Abhyankar, Eduardo Schwartz, David Lando, Denis Chaves, James Sefton, José Filipe Corrêa Guedes, Kent Daniel, Kristian Miltersen, Lasse Pederson, Nigel Barradale, Pasquale Della Corte, Richard Harris, Robert Kosowski, Soeren Hvidkjaer and numerous seminar participants at BI Norwegian School of Business, Copenhagen Business School, CUNEF, Imperial College, Nova School of Business and Economics, Universidad Carlos III and University of Exeter for their helpful comments and suggestions. A previous version of this paper circulated with the title “Managing the Risk of Momentum”. Santa-Clara is supported by a grant from the Fundação para a Ciência e Tecnologia (PTDC/EGE-GES/101414/2008).

†Nova School of Business and Economics. E-mail: pmsb@novasbe.pt.

‡Millennium Chair in Finance. Nova School of Business and Economics and NBER and CEPR. E-mail: psc@novasbe.pt.
1 Introduction

Momentum is a pervasive anomaly in asset prices. Jegadeesh and Titman (1993) find that previous winners in the US stock market outperform previous losers by as much as 1.49 percent a month. The Sharpe ratio of this strategy exceeds the Sharpe ratio of the market itself, as well as the size and value factors. Momentum returns are even more of a puzzle since they are negatively correlated to those of the market and value factors. Indeed, from 1927 to 2011, momentum had a monthly excess return of 1.75 percent controlling for the Fama-French factors.1 Moreover, momentum is not just a US stock market anomaly. Momentum has been documented in European equities, emerging markets, country stock indices, industry portfolios, currency markets, commodities and across asset classes.2 Grinblatt and Titman (1989,1993) find most mutual fund managers incorporate momentum of some sort in their investment decisions, so relative strength strategies are widespread among practitioners.

But the remarkable performance of momentum comes with occasional large crashes.3 In 1932, the winners-minus-losers strategy delivered a -91.59 percent return in just two months. In 2009, momentum experienced a crash of -73.42 percent in three months. Even the large returns of momentum do not compensate an investor with reasonable risk aversion for these sudden crashes that take decades to recover from.

The two most expressive momentum crashes occurred as the market rebounded following large previous declines. One explanation for this pattern is the time-varying systematic risk of the momentum strategy. Grundy and Martin (2001) show that momentum has significant negative beta following bear markets.4 They argue that hedging this time-varying market exposure produces stable momentum

---

1This result has led researchers to use momentum as an additional risk factor – Carhart (1997).
3Daniel and Moskowitz (2011)
4Following negative returns for the overall market, winners tend to be low-beta stocks and the reverse for losers. Therefore the winner-minus-losers strategy will have a negative beta.
returns but Daniel and Moskowitz (2012) show that using betas in real time does not avoid the crashes.

In this work we propose a different method to manage the risk of the momentum strategy. We estimate the risk of momentum by the realized variance of daily returns and find that it is highly predictable. An auto-regression of monthly realized variances yields an out-of-sample (OOS) R-square of 57.82 percent. This is 19.01 percentage points (p.p.) higher than a similar auto-regression for the variance of the market portfolio which is already famously predictable.\(^5\)

Furthermore, we find that high risk also forecasts lower momentum returns. In the years after a period of low realized volatility, the return of momentum is on average 24.35 percent. This contrasts with 0.96 percent in the years following a period of high realized volatility. The strategy delivers higher return when it is less risky, a clearly counter intuitive result that challenges any risk-based explanation for the strategy’s performance.

Managing the risk of momentum leads to substantial economic gains. We simply scale the long-short portfolio by its realized volatility in the previous 6 months, targeting a strategy with constant volatility.\(^6\) The Sharpe ratio improves from 0.53 for unmanaged momentum to 0.97 for its risk-managed version. But the most important benefit comes from a reduction in crash risk. The excess kurtosis drops from 18.24 to 2.68 and the left skew improves from -2.47 to -0.42. The minimum one-month return for raw momentum is -78.96 percent while for risk-managed momentum is -28.40 percent. The maximum drawdown of raw momentum is -96.69 percent versus -45.20 percent for its risk-managed version.

To assess the economic significance of our results, we evaluate the benefits of risk management for a risk-averse investor with a one-year horizon using a power utility function with Constant Relative Risk Aversion (CRRA) of four. Holding the market portfolio has an annual certainty equivalent of 0.14 percent to this investor. Adding momentum to the portfolio reduces the certainty equivalent to -5.46 percent. However, combining the market with risk-managed momentum achieves a certainty equivalent of 13.54 percent. We find that the main benefit of


\(^6\)This aproach only relies on past data and thus does not suffer from look-ahead bias. Scaling the portfolio to have constant volatility over time is actually a more natural way of implementing the strategy than having a constant amount long and short with varying volatility.
risk management comes in the form of lower crash risk. This alone provides a gain of 14.96 p.p. in annual certainty equivalent.

The performance of scaled momentum is robust in sub-samples and in international data. Using overlapping 10-year periods, we show that managing the risk of momentum not only avoids its worse crashes but it also improves the Sharpe ratio in 97.66 percent of the 10-year periods. Besides the US, risk-management also improves the Sharpe ratio of momentum in all the major markets we examine (France, Germany, Japan, and the UK).

One pertinent question is why managing risk with realized variances works while using time-varying betas does not. To answer this question we decompose the risk of momentum into market risk (from time-varying exposure to the market) and specific risk. We find that the market component is only 23 percent of total risk on average, so most of the risk of momentum is specific to the strategy. This specific risk is more persistent and predictable than the market component. The OOS R-square of predicting the specific component is 47 percent versus 21 percent for the systematic component. This is why hedging with time-varying betas fails: it focuses on the smaller and less predictable part of risk.

The research that is most closely related to ours is Grundy and Martin (2001) and Daniel and Moskowitz (2012). But their work studies the time-varying systematic risk of momentum, while we focus on momentum’s specific risk. Our results have the distinct advantage of offering investors using momentum strategies an effective way to manage risk without forward-looking bias. The resulting risk-managed strategy deepens the puzzle of momentum.

After the dismal performance of momentum in the last ten years, some could argue it is a dead anomaly. Our results indicate that momentum is not dead. It just so happens that the last ten years were rich in the kind of high-risk episodes that lead to bad momentum performance. Our paper is related to the recent literature that proposes alternative versions of momentum. Blitz, Huij and Martens (2011) show that sorting stocks according to their past residuals rather than gross returns produces a more stable version of momentum. Chaves (2012) shows that most of

---

7By momentum-specific risk we do not mean firm-specific risk or idiosyncratic risk. Momentum is a well diversified portfolio and all its risk is systematic. It just happens that a large portion of the risk of momentum is not explained by other risk factors.
the benefit in that method comes from using the market model in the regression and extends the evidence for residual momentum internationally.

Our work is also related to the literature on whether risk factors explain the risk of momentum (Griffin, Ji, and Martin (2003), Cooper, Gutierrez, and Hameed (2004), Fama and French (2012)).

This paper is structured as follows. Section 2 discusses the long-run properties of momentum returns and its exposure to crashes. Section 3 shows that momentum risk varies substantially over time in a highly predictable manner. Section 4 shows there is no risk-return trade-off for momentum. Actually we find the opposite of a trade-off. We analyze the implications of the predictability in momentum risk and return in Section 5. In Section 6 we use an utility based approach to decompose the gains of risk management according to the moments of returns. In Section 7 we decompose the risk of momentum and study the persistence of each of its components separately. In section 8 we check if our findings also hold internationally. In section 9 we assess the robustness of our findings across sub-samples. Finally, Section 10 presents our conclusions.

2 Momentum in the long run

We compare momentum to the Fama-French factors using a long sample of 85 years of monthly returns from July 1926 to December 2011. This is the same sample period as in Daniel and Moskowitz (2012).

Figure 1 presents the cumulative returns of each factor. As each factor consists of a long-short strategy, we construct the series of returns assuming the investor puts $1 in the risk-free asset at the beginning of the sample, buys $1 worth of the long portfolio and sells the same amount of the short portfolio. Then, in each subsequent month, the strategy fully reinvests the accumulated wealth in the risk-free asset, assuming a position of the same size in the long and short legs of the portfolio. The winners-minus-losers (WML) strategy offered an impressive performance for investors. One dollar fully reinvested in the momentum strategy grew to $68,741 by the end of the sample. This compares to the cumulative return

\(^8\)See the appendix for a description of the data.
of $2,136 from simply holding the market portfolio. Table 1 compares descriptive statistics for momentum in the long run with the Fama-French factors. Buying winners and shorting losers has provided large returns of 14.46 percent per year, with a Sharpe ratio higher than the market.

There would be nothing puzzling about momentum’s returns if they corresponded to a very high exposure to risk. However, running an OLS regression of the WML on the Fama-French factors gives (t-stats in parenthesis):

\[
\begin{align*}
  r_{WML,t} &= 1.752 - 0.378 r_{RMRF,t} - 0.249 r_{SMB,t} - 0.677 r_{HML,t} \\
  & (7.93) (-8.72) (-3.58) (-10.76)
\end{align*}
\]

so momentum provided abnormal returns of 1.75 percent per month after controlling for its negative exposure to the Fama and French (1992) risk factors. This amounts to a 21 percent per year abnormal return and the negative loadings on the risk factors imply momentum actually diversified risk in this extended sample.

The impressive excess returns of momentum, its high Sharpe ratio and negative relation to other risk factors, particularly the value premium, make it look like a free lunch to investors. But as Daniel and Moskowitz (2012) show there is a dark side to momentum. Its large gains come at the expense of a very high excess kurtosis of 18.24 combined with a pronounced left-skew of -2.47. These two features of the distribution of returns of the momentum strategy imply a very fat left tail – significant crash risk. Momentum returns can very rapidly turn into a free fall, wiping out decades of returns.

Figure 2 shows the performance of momentum in the two most turbulent decades for the strategy: the 1930’s and the 2000’s. In July and August 1932, momentum had a cumulative return of -91.59 percent. From March to May 2009, momentum had another large crash of -73.42 percent. These short periods had an enduring impact on cumulative returns. For example, someone investing one dollar in the WML strategy in July 1932 would only recover it in April 1963, 31 years later and with considerably less real value. This puts the risk to momentum investing in an adequate long-run perspective.

Both in 1932 and in 2009, the crashes happened as the market rebounded.
after experiencing large losses. This leads to the question of whether investors could have predicted the crashes in real time and hedge them away. Grundy and Martin (2001) show that momentum has a substantial time-varying loading on stock market risk. The strategy ranks stocks according to returns during a formation period, for example the previous 12 months. When the stock market performed well in the formation period, winners tend to be high-beta stocks and losers low-beta stocks. So the momentum strategy, by shorting losers to buy winners, has by construction a significant time-varying beta: positive after bull markets and negative after bear markets. They argue that hedging this time-varying risk produces stable returns, even in pre-WWII data, when momentum performed poorly. In particular, the hedging strategy would be long in the market portfolio whenever momentum has negative betas, hence mitigating the effects of rebounds following bear markets, which is when momentum experiences the worst returns. But the hedging strategy in Grundy and Martin (2001) uses forward looking betas, estimated with information investors did not have in real time. Using betas estimated solely on ex-ante information does not avoid the crashes and portfolios hedged in real time often perform even worse than the original momentum strategy (Daniel and Moskowitz (2012), Barroso (2012)).

3 The time-varying risk of momentum

One possible cause for excess kurtosis is time-varying risk. The very high excess kurtosis of 18.24 of the momentum strategy (more than twice the market portfolio) leads us to study the dynamics of its risk and compare it with the risk of the market (RMRF), value (HML) and size (SMB) risk factors.

For each month, we compute the realized variance $RV_t$ from daily returns in the previous 21 sessions. Let $\{r_d\}_{d=1}^{D_t}$ be the daily returns and $\{d_t\}_{t=1}^{T_t}$ the time series of the dates of the last trading sessions of each month. Then the realized variance of factor $i$ in month $t$ is:\footnote{Daniel and Moskowitz (2011) argue this is due to the option-like payoffs of distressed firms in bear markets.}

\footnote{See, for example, Engle (1982) and Bollerslev (1987).}

\footnote{Correcting for serial correlation of daily returns does not change the results significantly.}
Figure 3 shows the monthly realized volatility of momentum. This varies substantially over time, from a minimum of 3.04 percent (annualized) to a maximum of 127.87 percent.

Table 2 shows the results of AR (1) regressions of the realized variances of the WML, RMRF, SMB and HML:

\[ RV_{i,t} = \sum_{j=0}^{20} r_{i,d_t-j}^2 \]  

(2)

\[ RV_{i,t} = \alpha + \rho RV_{i,t-1} + \varepsilon_t \]  

(3)

Panel A presents the results for RMRF and WML, for which we have daily data available from 1927:03 to 2011:12. Panel B adds the results for HML and SMB, for which daily data is available only from 1963:07 onwards.

Momentum returns are the most volatile. From 1927:03 to 2011:12, the average realized volatility of momentum is 15.03, more than the 12.81 of the market portfolio. For 1963:07 onwards, the average realized volatility of momentum is 16.40, also the highest when compared to RMRF, SMB and HML.

In the full sample period, the standard deviation of monthly realized volatilities is higher for momentum (12.26) than the market (7.82). Panel B confirms this result compared to the other factors in the 1963:07 onwards sample. So the risk of momentum is the most variable.

The risk of momentum is also the most persistent. The \( AR(1) \) coefficient of the realized variance of momentum in the 1963:07 sample is 0.77, which is 0.19 more than for the market and higher than the estimates for SMB and HML.

To check the out-of-sample (OOS) predictability of risk, we use a training sample of 240 months to run an initial AR(1) and then use the estimated coefficients and last observation of realized variance to forecast the realized variance in the following month. Then each month we use an expanding window of observations to produce OOS forecasts and compare these to the accuracy of the historical mean \( \bar{RV}_{i,t} \). As a measure of goodness of fit we estimate the OOS R-square as:
\[ R^2_{i,OOS} = 1 - \frac{\sum_{t=S}^{T-1} (\hat{\alpha}_t + \hat{\rho}_t RV_{i,t} - RV_{i,t+1})^2}{\sum_{t=S}^{T-1} (RV_{i,t} - RV_{i,t+1})^2} \] (4)

where \( S \) is the initial training sample, \( \hat{\alpha}_t, \hat{\rho}_t \) and \( RV_{i,t} \) are estimated with information available only up to time \( t \).

The last column of table 2 shows the OOS R-squares of each auto-regression. The AR(1) of the realized variance of momentum has an OOS R-square of 57.82 percent (full sample), which is 50 percent more than the market. For the period from 1963:07 to 2011:12, the OOS predictability of momentum risk is twice that of the market. Hence more than half of the risk of momentum is predictable, the highest level among risk factors.

### 4 No risk-return trade-off in momentum

Predictable risk does not necessarily imply the possibility of achieving economic gains by managing risk. If higher risk also forecasts higher returns, the implicit trade-off could leave investors indifferent. But in fact, we find exactly the opposite. Higher risk in the momentum factor forecasts lower expected returns going forward.

A predictive regression of the monthly returns of momentum on lagged 6-month realized variance holds (t-stats in parenthesis): \[ r_{WML,t+1} = 1.79_{(6.46)} -1.51_{(-4.63)} RV_{WML(-6),t} \]

The monthly returns are from 1927:03 to 2011:12. The regression has an R-square of 2.06 percent, which is high for a predictive regression at monthly frequency. Moreover, the OOS R-square, considering a training sample of 240 months, is of 1.85 percent. So, unlike most predictive regressions for the market return (Goyal and Welch (2008)), the regression is robust out of sample.\(^{12}\)

Figure 4 illustrates the potential of realized variance of momentum to condition exposure to the factor. We sort the months into quintiles according to the level

\(^{12}\)Examining the two legs separately, we find the variance of momentum forecasts both higher returns for the losers portfolio and lower returns for the winners portfolio. But predictability is much stronger for the winner-minus-losers portfolio.
of realized variance in the previous 6 months for each factor: the market and momentum. Quintile 1 is the set of months with lowest risk and quintile 5 is the one with highest risk. Then we report, for each factor, the average realized volatility, return and Sharpe ratio in the following 12 months.

In general, higher risk in the recent past forecasts higher risk going forward. This is true both for the market and the momentum factors, but more so in the case of momentum.

For the market there is no obvious trade-off between risk and return. This illustrates the well-known difficulty in estimating this relation. But this is not true for momentum. The winners-minus-losers portfolio has a very clear negative relation between risk and return. In the years following calm periods (in quintile 1), the average annual return of the momentum strategy is 24.35%. By contrast, in the years following a turbulent period (in quintile 5), the average return of the momentum strategy is only 0.96%.

This negative relation between risk and return adds a new layer to the momentum puzzle. Higher risk forecasts both higher risk and lower returns. As a consequence the Sharpe ratio of the momentum factor changes considerably conditional on its previous risk. In the years after a calm period, the Sharpe ratio is 1.72 on average. By contrast, after a turbulent period the Sharpe ratio is only 0.28 on average.

In the next section, we explore the combined potential of this predictability in returns with the predictability of risk and show their usefulness to manage the exposure to momentum.

5 Risk-managed momentum

We use an estimate of momentum risk to scale the exposure to the strategy in order to have constant risk over time. Volatility-scaling has been used in the time series momentum literature (Moskowitz, Ooi, and Pederson (2012) and Baltas and Kosowski (2013)). But there it serves a different purpose: use the asset-specific volatility to prevent the results from being dominated by high-volatility assets only. We do not consider asset-specific volatilities and actually show that the persistence in risk of the winners-minus-losers strategy is more interesting than that of a long-only portfolio.
forecast $\hat{\sigma}_t^2$ from daily returns in the previous 6 months. Let \{$r_{WML,t}$\}$^T_{t=1}$ be the monthly returns of momentum and \{$r_{WML,d}, d=1, \ldots, T$\}, \{$d_t\}$ be, as above, the daily returns and the time series of the dates of the last trading sessions of each month.

The variance forecast is:

$$\hat{\sigma}_{WML,t}^2 = 21 \sum_{j=0}^{125} r_{WML,d, t-1-j}^2 / 126$$

(5)

As $WML$ is a zero-investment and self-financing strategy we can scale it without constraints. We use the forecasted variance to scale the returns:

$$r_{WML^*, t} = \frac{\sigma_{\text{target}}}{\hat{\sigma}_t} r_{WML, t}$$

(6)

where $r_{WML, t}$ is the unscaled or plain momentum, $r_{WML^*, t}$ is the scaled or risk-managed momentum, and $\sigma_{\text{target}}$ is a constant corresponding to the target level of volatility. Scaling corresponds to having a weight in the long and short legs that is different from one and varies over time, but it keeping the strategy still self-financing. We pick a target corresponding to an annualized volatility of 12 percent.

Figure 5 shows the cumulative returns of risk-managed momentum compared to plain momentum. The risk-managed momentum strategy achieves a higher cumulative return with less risk. So there are economic gains to risk management of momentum. The scaled strategy benefits from the large momentum returns when it performs well and effectively shuts off in turbulent times, thus mitigating momentum crashes. As a result, one dollar invested in risk-managed momentum grows to $6,140,075 by the end of the sample, nearly 90 times more than the plain momentum strategy. Also, the risk-managed strategy no longer has variable and of assets, such as the market.

15 We also used one-month and three-month realized variances as well as exponentially-weighted moving average (EWMA) with half-lives of 1, 3 and 6 months. All work well with nearly identical results.

16 The annualized standard deviation from monthly returns will be higher than 12% as volatilities at daily frequency are not directly comparable to those at lower frequencies due to the small positive autocorrelation of daily returns.

17 This difference in cumulative returns is fundamentally due to risk management successfully avoiding the two momentum crashes. But in the post-war period from 1946 to 2007, not including the crashes, the Sharpe ratio of momentum was 0.86, versus 1.15 for risk-managed momentum.
persistent risk, so risk management does indeed work.\textsuperscript{18}

Table 3 provides a summary of the economic performance of $WML^{*}$ in 1927-
2011. The risk-managed strategy has a higher average return, with a gain of 2.04
percentage points per year, with substantially less standard deviation (less 10.58
percentage points per year). As a result, the Sharpe ratio of the risk-managed
strategy almost doubles from 0.53 to 0.97. The most important gains of risk-
management show up in the improvement in the higher order moments. Managing
the risk of momentum lowers the excess kurtosis from a very high value of 18.24 to
just 2.68 and reduces the left skew from -2.47 to -0.42. This practically eliminates
the crash risk of momentum. Figure 6 shows the density function of momentum
and its risk-managed version. Momentum has a very long left tail which is much
reduced in its risk-managed version.

The benefits of risk-management are especially important in turbulent times.
Figure 7 shows the performance of risk-managed momentum in the decades with
the most impressive crashes. The scaled momentum manages to preserve the
investment in the 1930’s. This compares very favorably to the pure momentum
strategy which loses 90 percent in the same period. In the 2000’s simple momentum
loses 28 percent of wealth, because of the crash in 2009. Risk-managed momentum
ends the decade up 88 percent as it not only avoids the crash but also captures
part of the positive returns of 2007-2008.

Figure 8 shows the weights of the scaled momentum strategy over time – inter-
preted as the dollar amount in the long or short leg. These range between the
values of 0.13 and 2.00, reaching the most significant lows in the early 1930’s, in
2000-02, and in 2008-09. On average, the weight is 0.90, slightly less than full
exposure to momentum. As these weights depend only on ex-ante information
this strategy could actually be implemented in real time. Actually, running a
long-short strategy to have constant volatility is closer to what real investors (like
hedge funds) try to do than keeping a constant amount invested in the long and
short legs of the strategy.

\textsuperscript{18}The AR(1) coefficient of monthly squared returns is only 0.14 for the scaled momentum versus
0.40 for the original momentum. Besides, the auto-correlation of raw momentum is significant up
to 15 lags while only 1 lag for risk-managed momentum. So persistence in risk is much smaller
for the risk-managed strategy.
6 Economic significance: An investor perspective

Raw momentum offers a trade-off between an appealing Sharpe ratio, obtained from the first two moments of its distribution, and less appealing higher order moments, such as high kurtosis and left skewness. An economic criterion is needed to assess whether this trade-off is interesting. Risk management offers improvements to momentum across the board, higher expected returns, lower standard deviation and crash risk. Still it is pertinent to evaluate the relative economic importance of each of these contributions.

We use an utility-based approach to discuss the appeal of momentum to a representative investor. We adopt the power utility function as it has the advantage of taking into consideration higher order moments instead of focusing merely on the mean and standard deviation of returns. This is particularly important since momentum has a distribution far from normal. The utility of returns is:

\[ U(r) = \frac{(1 + r)^{1-\gamma}}{1 - \gamma} \]  

where \( \gamma \) is the Constant Coefficient of Relative risk aversion (CRRA). Bliss and Panigirtzoglou (2004) estimate \( \gamma \) empirically from risk-aversion implicit in one-month options on the S&P and the FTSE and find a value very close to 4. This is a more plausible value for CRRA than previous estimates in the equity premium puzzle literature using utility over consumption. So we adopt this value for \( \gamma \). We obtain the certainty equivalent from the utility of returns:

\[ CE(r) = \left\{ (1 - \gamma)E[U(r)] \right\}^{\frac{1}{1-\gamma}} - 1 \]  

This states the welfare a series of returns offers the investor in terms of an equivalent risk-free annual return, expressed in a convenient unit of percentage points per year.

For an economic measure of the importance of the mean return, variance and higher order moments, we use a Taylor series approximation to expected utility around the mean:
\[ E[U(r)] = U(E(r)) + \frac{1}{2} U''(E(r)) E(r - E(r))^2 + \phi_3(r) \] \hspace{1cm} (9)

where \( \phi_3 \) is the remainder corresponding to the utility from moments with order greater than 2. From this we obtain the certainty equivalent due to each moment:

\[ CE(\mu_1) = \left\{ (1 - \gamma)U(E(r)) \right\}^{\frac{1}{1-\gamma}} - 1 \] \hspace{1cm} (10)

\[ CE(\mu_2) = \left\{ (1 - \gamma)[U(E(r)) + \frac{1}{2} U''(E(r)) E(r - E(r))^2] \right\}^{\frac{1}{1-\gamma}} - CE(\mu_1) - 1 \] \hspace{1cm} (11)

\[ CE(\mu_{h>2}) = CE(r) - CE(\mu_1) - CE(\mu_2) \] \hspace{1cm} (12)

We compute the certainty equivalent from annual overlapping returns. A one-year horizon captures the occasional large drawdowns of momentum documented in Section 2.

Table 4 shows the decomposition of the certainty equivalent for the representative investor holding 100% of the wealth in the market portfolio. It also assesses whether it is optimal to deviate from the market portfolio to include some weight in a long-short strategy such as momentum.

The first row shows the results for holding only the market portfolio. The mean return had a positive contribution for the certainty equivalent of 11.72 percent per year, but the variance of returns reduces this by 7.39 percentage points. Higher order moments diminish the certainty equivalent by a further 4.18 percentage points. As a result the certainty equivalent of the market portfolio was only 0.14 percent per year.

Adding momentum to the market portfolio increases returns. As a result, the certainty equivalent of the mean return increases from 11.72 percent per year to 28.51 percent. The higher standard deviation partially offsets this gain by reducing the certainty equivalent by 6.51 percentage points. Still, looking only at the first two moments of the combined portfolio leads to the conclusion that the investor is better off including momentum.
But the increase in higher order risk – the momentum crashes – reduces the certainty equivalent by 15.89 percentage points per year. As a result, including momentum actually reduces the economic performance of the market portfolio. The certainty equivalent of the market plus momentum is -5.46 percent per year versus 0.14 percent of the market only. So, in spite of the impressive cumulative returns of momentum in the long-run, crash risk is so high that a reasonably risk-averse investor would rather just hold the market portfolio.

This illustrates with an economic measure how far the distribution of momentum is from normality. Indeed, momentum has a distribution with many small gains and few, but extreme, large losses. Taking this into account, the momentum puzzle of Jegadeesh and Titman (1993) is substantially diminished.

In contrast, risk-managed momentum produces large economic gains. These come from higher returns when compared to the market (a 19.92 percentage point gain) and less crash risk than the market with plain momentum (a 14.96 percentage point gain). As a result, the annual certainty equivalent of the market with risk-managed momentum is 13.54 percent, which compares very favorably to the 0.14 percent of the market alone and even more so with the -5.46 percent of the market combined with raw momentum.

7 Anatomy of momentum risk

A well documented result in the momentum literature is that momentum has time-varying market risk (Grundy and Martin (2001)). This is an intuitive finding since after bear markets winners are low-beta stocks and the losers have high betas. But Daniel and Moskowitz (2012) show that using betas to hedge risk in real time does not work. This contrasts with our finding that the risk of momentum is highly predictable and managing it offers strong gains. Why is scaling with forecasted variances so different from hedging with market betas? We show it is because time-varying betas are not the main source of predictability in momentum risk.

We use the market model to decompose the risk of momentum into market and specific risk:

$$RV_{wm,l,t} = \beta_t^2 RV_{mr,f,t} + \sigma_{e,l}^2$$  \hspace{1cm} (13)
The realized variances and betas are estimated with 6-months of daily returns. On average, the market component $\beta^2_t RV_{rmrf,t}$ accounts for only 23 percent of the total risk of momentum. Almost 80 percent of the momentum risk is specific to the strategy. Also, the different components do not have the same degree of predictability. Table 5 shows the results of an AR (1) on each component of risk.

Either in-sample or out-of-sample (OOS), $\beta^2_t$ is the least predictable component of momentum risk. Its OOS R-square is only 5 percent. The realized variance of the market also has a small OOS R-square of 7 percent. When combined, both elements form the market risk component and show more predictability with an OOS R-square of 21 percent, but still less than the realized variance of momentum with an OOS R-square of 44 percent. The most predictable component of momentum variance is the specific risk with an OOS R-square of 47 percent, more than double the predictability of the part due to the market.

Hedging the market risk alone, as in Daniel and Moskowitz (2012) fails because most of the risk is left out.\footnote{One alternative way to decompose the risk of the winner-minus-losers portfolio is to consider the variance of the long and short leg separately and also their covariance. We examine this alternative decomposition and find the predictability of the risk of momentum cannot be fully attributed to just one of its components. They all show predictability and contribute to the end result. We omit the results for the sake of brevity.}

### 8 International Evidence

Recently, Chaves (2012) examines the properties of momentum in an international sample of 21 countries. We use his data set, constructed from Datastream stock-level data, to check if our results also hold in international equity markets. Chaves (2012) requires at least 50 listed stocks in each country, selects only those in the top half according to market capitalization and further requires that they comprise at least 90% of the total market capitalization of each country. This aims to ensure that the stocks considered are those of the largest, most representative and liquid shares in each market. Then he sorts stocks into quintiles according to previous returns from month t-12 to t-2 and computes winners-minus-losers returns at daily and monthly frequencies for each country. These are defined as the difference...
between the return of winner quintile and the loser quintile.\textsuperscript{20} There are only four countries that consistently satisfy the minimum data requirements: France, Germany, Japan and the United Kingdom. We choose to only report results for these four countries for which data is more reliable and therefore avoid possible issues with missing observations and outliers.\textsuperscript{21}

Table 6 shows the results of the risk-managed momentum in the four countries considered. The risk-managed strategy uses, as above, the realized volatility in the previous 6 months to target a constant volatility of 12 percent (annualized).

Managing the risk of momentum improves both the Sharpe ratio and the certainty equivalent in all four countries. Notably in Japan, where the momentum strategy usually fails (Asness (2011), Chaves (2012)) we find that managing the risk of momentum triples the Sharpe ratio from an insignificant 0.08 to (a still modest) 0.24. Risk management improves the Sharpe ratio up to 0.68 (in the UK) and the annual certainty equivalent up to 22.49 percentage points (also in the UK).

As such we conclude that the benefits of managing the risk of momentum are pervasive in international data. Since we did not have this data at the time of writing the first version of the paper, we see this as true out-of-sample confirmation of our initial results.

9 Robustness checks

As figure 2 shows, managing the risk of momentum makes a crucial difference to investors at times of greater uncertainty. This begs the reverse question of whether there are any benefits of risk-management in less turbulent times too. There are two crashes of a very large magnitude in our sample, the first one in 1932 and the second in 2009. Therefore it is pertinent to ask to what extent are our entire results totally driven by these two singular occurrences.

To address this issue we examine the performance of both momentum and risk-managed momentum in the relatively benign period of 1945:01 to 2005:12. This

\textsuperscript{20}We thank Denis Chaves for letting us use his data. We refer to his paper for a more detailed description of the international data used.

\textsuperscript{21}Nevertheless, we check if managing the risk of momentum improves its performance in the other 17 countries with less reliable data. We find that it does improve the Sharpe ratio in all of those countries.
roughly corresponds to the post-war period up to the years preceding the Great Recession. Figure 9 shows the cumulative returns of both investment strategies. In this period the performance of the raw momentum strategy is much more impressive than in the more extended sample we discuss in section 2. The Sharpe ratio of the WML portfolio is 0.86, the excess kurtosis 6.00 and the skewness -0.91. This makes it a formidable benchmark to outperform. But our risk-managed momentum still has better performance in this period, achieving a higher cumulative return with less risk. Its Sharpe ratio in this period is 1.16, the excess kurtosis of 1.20 is much smaller than for raw WML and the skewness of -0.17 is closer to zero.

We also check if risk-management produces robust economic gains across subsamples. We consider the performance of both strategies in rolling (overlapping) one-year, 3-year, 5-year and 10-year horizons and compute the Sharpe ratios and certainty equivalents for each period.

The benefits of managing risk are stronger the longer the horizon considered. Table 7 shows that risk management increases the Sharpe ratio of 60.38 percent of one-year periods and 97.66 percent of 10-year periods. The results for the certainty equivalent confirm this pattern. On average, for a 5-year horizon, managing the risk of momentum produces a 10.66 percentage points gain in annual certainty equivalent and leads to outperformance in 85.61 percent of the periods. So, considering a long enough investment horizon, risk management produces large and consistent economic gains.

Overall, we conclude the results are robust across subsamples and are not driven by just rare events in the 1930’s and in the 2000’s.

Another issue is to what extent our results overlap with other research on the predictability of momentum’s risk and return. Most of that literature has focused on the time-varying market risk of momentum. Grundy and Martin (2001) show how the beta of the momentum strategy changes over time with lagged market returns. Cooper, Gutierrez, and Hameed (2004) show that momentum’s expected returns depend on the state of the market. Daniel and Moskowitz (2012) show that momentum crashes follow a pattern, occurring during reversals after a bear market. We compare the predictive power of our conditional variable – realized volatility of momentum - with a bear-market state variable. We find that the realized volatility of momentum has an informational content – forecasting both
returns and risk – that is much greater and more robust than the bear-market indicator.\textsuperscript{22}

10 Conclusion

Unconditional momentum has a distribution that is far from normal, with huge crash risk. We find that taking this crash risk into consideration, momentum is not appealing for a risk-averse investor.

However, we find the risk of momentum is highly predictable. Managing this risk eliminates exposure to crashes and increases the Sharpe ratio of the strategy substantially. This presents a new challenge to any theory attempting to explain momentum.

Our results are confirmed with international evidence and robust across sub-samples.

References


\textsuperscript{22}We omit those results for the sake of brevity.


Appendix: Data sources

We obtain daily and monthly returns for the market portfolio, the high-minus-low, the small-minus-big, the ten momentum-sorted portfolios and the risk-free (one-month Treasury-bill return) from Kenneth French’s data library. The monthly data is from July 1926 to December 2011 and the daily data is from July 1963 to December 2011.

For the period from July 1926 to June 1963, we use daily excess returns on the market portfolio (the value-weighted return of all firms on NYSE, AMEX and Nasdaq) from the Center for Research in Security Prices (CRSP). We also have daily returns for ten value-weighted portfolios sorted on previous momentum from Daniel and Moskowitz (2012). This allows us to work with a long sample of daily returns for the winner-minus-losers (WML) strategy from August 1926 to December 2011. We use these daily returns to calculate the realized variances in the previous 21, 63 and 126 sessions at the end of each month.

For the momentum portfolios, all stocks in the NYSE, AMEX and Nasdaq universe are ranked according to returns from month t-12 to t-2, then classified into deciles according to NYSE cutoffs. So there is an equal number of NYSE firms in each bin. The WML strategy consists on shorting the lowest (loser) decile and a long position in the highest (winner) decile. Individual firms are value weighted in each decile. Following the convention in the literature, the formation period for month t excludes the returns in the preceding month. See Daniel and Moskowitz (2012) for a more detailed description of how they build momentum portfolios. The procedures (and results) are very similar to those of the Fama-French momentum portfolios for the 1963:07-2011:12 sample.
### Table 1.
The long-run performance of momentum compared to the Fama-French factors. All statistics computed with monthly returns. Max and Min are the maximum and minimum one-month returns observed in the sample. Mean is the average excess return (annualized) and ‘STD’ is the (annualized) standard deviation of each factor. ‘KURT’ stands for excess kurtosis and ‘SR’ for (annualized) Sharpe ratio. The sample returns are from 1927:03 to 2011:12.

<table>
<thead>
<tr>
<th></th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>STD</th>
<th>KURT</th>
<th>SKEW</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMRF</td>
<td>38.27</td>
<td>-29.04</td>
<td>7.33</td>
<td>18.96</td>
<td>7.35</td>
<td>0.17</td>
<td>0.39</td>
</tr>
<tr>
<td>SMB</td>
<td>39.04</td>
<td>-16.62</td>
<td>2.99</td>
<td>11.52</td>
<td>21.99</td>
<td>2.17</td>
<td>0.26</td>
</tr>
<tr>
<td>HML</td>
<td>35.48</td>
<td>-13.45</td>
<td>4.50</td>
<td>12.38</td>
<td>15.63</td>
<td>1.84</td>
<td>0.36</td>
</tr>
<tr>
<td>WML</td>
<td>26.18</td>
<td>-78.96</td>
<td>14.46</td>
<td>27.53</td>
<td>18.24</td>
<td>-2.47</td>
<td>0.53</td>
</tr>
</tbody>
</table>
Panel A: 1927:03 to 2011:12

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>t-stat</th>
<th>( \rho )</th>
<th>t-stat</th>
<th>( R^2 )</th>
<th>OOS ( R^2 )</th>
<th>( \sigma )</th>
<th>( \sigma_\alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMRF</td>
<td>0.0010</td>
<td>6.86</td>
<td>0.60</td>
<td>23.92</td>
<td>36.03</td>
<td>38.81</td>
<td>12.81</td>
<td>7.82</td>
</tr>
<tr>
<td>WML</td>
<td>0.0012</td>
<td>5.21</td>
<td>0.70</td>
<td>31.31</td>
<td>49.10</td>
<td>57.82</td>
<td>15.03</td>
<td>12.26</td>
</tr>
</tbody>
</table>

Panel B: 1963:07 to 2011:12

<table>
<thead>
<tr>
<th></th>
<th>( \alpha )</th>
<th>t-stat</th>
<th>( \rho )</th>
<th>t-stat</th>
<th>( R^2 )</th>
<th>OOS ( R^2 )</th>
<th>( \sigma )</th>
<th>( \sigma_\alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMRF</td>
<td>0.0009</td>
<td>5.65</td>
<td>0.58</td>
<td>17.10</td>
<td>33.55</td>
<td>25.46</td>
<td>13.76</td>
<td>8.48</td>
</tr>
<tr>
<td>SMB</td>
<td>0.0004</td>
<td>8.01</td>
<td>0.33</td>
<td>8.32</td>
<td>10.68</td>
<td>-8.41</td>
<td>7.36</td>
<td>3.87</td>
</tr>
<tr>
<td>HML</td>
<td>0.0001</td>
<td>4.88</td>
<td>0.73</td>
<td>25.84</td>
<td>53.55</td>
<td>53.37</td>
<td>6.68</td>
<td>4.29</td>
</tr>
<tr>
<td>WML</td>
<td>0.0009</td>
<td>3.00</td>
<td>0.77</td>
<td>29.29</td>
<td>59.71</td>
<td>55.26</td>
<td>16.40</td>
<td>13.77</td>
</tr>
</tbody>
</table>

Table 2. AR (1) of one-month realized variances. The realized variances are the sum of squared daily returns in each month. The AR (1) regresses the non-overlapping realized variance of each month on its own lagged value and a constant. The OOS R-square uses the first 240 months to run an initial regression so producing an OOS forecast. Then uses an expanding window of observations till the end of the sample. In panel A the sample period is from 1927:03 to 2011:12. In panel B we repeat the regressions for RMRF and WML and add the same information for the HML and SMB. The last two columns show, respectively, the average realized volatility and its standard deviation.

<table>
<thead>
<tr>
<th></th>
<th>Max</th>
<th>Min</th>
<th>Mean</th>
<th>STD</th>
<th>KURT</th>
<th>SKEW</th>
<th>SR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMRF</td>
<td>25.93</td>
<td>-3.96</td>
<td>13.46</td>
<td>23.78</td>
<td>18.24</td>
<td>-2.47</td>
<td>0.53</td>
</tr>
<tr>
<td>SMB</td>
<td>21.95</td>
<td>-28.40</td>
<td>16.50</td>
<td>16.95</td>
<td>2.68</td>
<td>-0.42</td>
<td>0.97</td>
</tr>
</tbody>
</table>

Table 3. Momentum and the economic gains from scaling. The first row presents as a benchmark the economic performance of plain momentum from 1927:03 to 2011:12. The second row presents the performance of risk-managed momentum. The risk-managed momentum uses the realized variance in the previous 6 months to scale the exposure to momentum.
Table 4. The economic performance of momentum for a representative investor. The first row shows the performance of the market portfolio. The second row combines the market portfolio with momentum and the third one with scaled momentum. The first column shows the certainty equivalent of the mean return of each strategy. The second and third columns present the contribution to the certainty equivalent of standard deviation and higher moments, respectively. The last column shows the certainty equivalent obtained from annual non-overlapping returns. The returns are from 1927:03 to 2011:12. The decomposition uses a Taylor expansion of the utility function around the mean return of the portfolio with a CRRA of 4.

<table>
<thead>
<tr>
<th></th>
<th>$CE(\mu_1)$</th>
<th>$CE(\mu_2)$</th>
<th>$CE(\mu_{i&gt;2})$</th>
<th>$CE(r)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RM</td>
<td>11.72</td>
<td>-7.39</td>
<td>-4.18</td>
<td>0.14</td>
</tr>
<tr>
<td>RM+WML</td>
<td>28.51</td>
<td>-13.90</td>
<td>-20.07</td>
<td>-5.46</td>
</tr>
<tr>
<td>RM+WML*</td>
<td>31.64</td>
<td>-13.00</td>
<td>-5.11</td>
<td>13.54</td>
</tr>
</tbody>
</table>

Table 5. Decomposition of the risk of momentum. Each row shows the results of an AR (1) for 6-month, non-overlapping periods. The first row is for the realized variance of the WML and the second one the realized variance of the market. The third row is squared beta, estimated as a simple regression of 126 daily returns of the WML on RMRF. The fourth row is the systematic component of momentum risk and the last row the specific component. The OOS R-squares use an expanding window of observations after an initial in-sample period of 20 years.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma^2_{wml}$</th>
<th>$\sigma^2_{mr rf}$</th>
<th>$\beta^2$</th>
<th>$\beta^2 \sigma^2_{r mrf}$</th>
<th>$\sigma^2_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.0012</td>
<td>0.0012</td>
<td>0.3544</td>
<td>0.0007</td>
<td>0.0007</td>
</tr>
<tr>
<td>t-stat</td>
<td>2.59</td>
<td>4.29</td>
<td>6.05</td>
<td>2.73</td>
<td>2.69</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.70</td>
<td>0.50</td>
<td>0.21</td>
<td>0.47</td>
<td>0.72</td>
</tr>
<tr>
<td>t-stat</td>
<td>12.58</td>
<td>7.37</td>
<td>2.83</td>
<td>6.80</td>
<td>13.51</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.49</td>
<td>0.25</td>
<td>0.05</td>
<td>0.22</td>
<td>0.52</td>
</tr>
<tr>
<td>$R^2_{OOS}$</td>
<td>0.44</td>
<td>0.07</td>
<td>0.05</td>
<td>0.21</td>
<td>0.47</td>
</tr>
</tbody>
</table>
Table 6. The Sharpe ratio and the certainty equivalent of plain momentum and scaled momentum (all measures are annualized). Original data from Datastream. The returns for country-specific momentum portfolios are from Chaves (2012). The scaled momentum uses the realized volatility in the previous 6-months, obtained from daily data. The certainty equivalent uses annual overlapping returns and considers a CRRA of 4. The certainty equivalent assumes an investment in the domestic risk-free rate asset of the US for all countries plus the respective long-short stock portfolio. The returns are from 1980:07 to 2011:10.

<table>
<thead>
<tr>
<th>Country</th>
<th>SR</th>
<th>ΔSR</th>
<th>CE</th>
<th>ΔCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WML</td>
<td>WML*</td>
<td>WML</td>
<td>WML*</td>
</tr>
<tr>
<td>FRANCE</td>
<td>0.67</td>
<td>1.03</td>
<td>0.36</td>
<td>2.50%</td>
</tr>
<tr>
<td>GERMANY</td>
<td>1.02</td>
<td>1.39</td>
<td>0.38</td>
<td>11.32%</td>
</tr>
<tr>
<td>JAPAN</td>
<td>0.08</td>
<td>0.24</td>
<td>0.15</td>
<td>-5.70%</td>
</tr>
<tr>
<td>UNITED KINGDOM</td>
<td>1.09</td>
<td>1.77</td>
<td>0.68</td>
<td>5.27%</td>
</tr>
</tbody>
</table>

Table 7. Comparison of rolling-period Sharpe ratios and certainty equivalents of raw momentum and risk-managed momentum (denoted with a star). We compute the monthly returns of both strategies for rolling, overlapping, intervals of one, three, five and ten years. From these, we obtain the certainty equivalent (assuming a CRRA of 4) and the Sharpe ratio for each interval. The first column reports the average difference in certainty equivalent of the risk-managed strategy over raw momentum. The second column shows the percentage of intervals for which the risk-managed strategy outperforms the plain one. Columns 3 and 4 show the same information for the Sharpe ratio. The returns are from 1927:03 to 2011:12.

<table>
<thead>
<tr>
<th>Periods</th>
<th>CE*-CE</th>
<th>CE*&gt;CE (%)</th>
<th>SR*-SR</th>
<th>SR*&gt;SR (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year</td>
<td>3.09</td>
<td>62.96</td>
<td>0.06</td>
<td>60.38</td>
</tr>
<tr>
<td>3 years</td>
<td>7.95</td>
<td>76.30</td>
<td>0.15</td>
<td>79.15</td>
</tr>
<tr>
<td>5 years</td>
<td>10.66</td>
<td>85.61</td>
<td>0.17</td>
<td>83.42</td>
</tr>
<tr>
<td>10 years</td>
<td>11.12</td>
<td>91.21</td>
<td>0.18</td>
<td>97.66</td>
</tr>
</tbody>
</table>
Figure 1. The long-run cumulative returns of momentum compared to the Fama-French factors. Each strategy consists on investing $1 at the beginning of the sample in the risk-free rate and combine it with the respective long-short portfolio. The proceeds are fully reinvested till the end of the sample. On the right is the terminal value of each strategy.
Figure 2. Momentum crashes. The figure plots the cumulative return and terminal value of the momentum and market portfolio strategies in its two most turbulent periods: the 1930’s and the 2000’s.
Figure 3. The realized volatility of momentum obtained from daily returns in each month from 1927:03 to 2011:12.
Figure 4. The performance of the market factor (‘RMRF’) and momentum (‘WML’) conditional on realized volatility in the previous 6 months. Returns are sorted into quintiles according with volatility in the previous 6 months for each factor. The figure presents the following year volatility (computed from daily returns), the cumulative return in percentage points and the Sharpe Ratio. The returns are from 1927:03 to 2011:12.
Figure 5. The long-run performance of risk-managed momentum. The risk-managed momentum (WML*) scales the exposure to momentum using the realized variance in the previous 6-months. In the beginning of the sample the strategy invests $1 in the risk-free asset and combines it with the long-short portfolio. The proceeds are fully reinvested till the end of the sample. On the right is the terminal value of the strategy.
Figure 6. The density of plain momentum (WML) and risk-managed momentum (WML*). The risk-managed momentum uses the realized variance in the previous 6 months to scale the exposure to momentum. The returns are from 1927:03 to 2011:12.
Figure 7. The benefits of risk-management in the 1930’s and the 00’s. The risk-managed momentum (WML*) uses the realized variance in the previous 6 months to scale the exposure to momentum.
Figure 8. Weights of the scaled momentum. The risk-managed momentum uses the realized variance in the previous 6 months to scale the exposure to momentum.
Figure 9. Cumulative return of momentum and risk-managed momentum from 1945:01 to 2005:12. The risk-managed momentum uses the realized volatility of the momentum strategy in the previous 6 months to scale the exposure and achieve a target constant volatility of 12 percent (annualized).