Journal of Financial Economics 116 (2015) 111–120

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/jfec

Momentum has its moments ☆

Pedro Barroso [a,*], Pedro Santa-Clara [b,c,d]

[a] University of New South Wales, Australia
[b] Nova School of Business and Economics, Portugal
[c] NBER, United States
[d] CEPR, United Kingdom

ARTICLE INFO

Article history:
Received 14 May 2013
Received in revised form
9 December 2013
Accepted 20 January 2014
Available online 29 November 2014

JEL classification:
G11
G12
G17

Keywords:
Anomalies
Momentum
Time-varying risk
Transaction costs of momentum

ABSTRACT

Compared with 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 who dislike negative skewness and kurtosis. We find that the risk of momentum is highly variable over time and predictable. Managing this risk virtually eliminates crashes and nearly doubles the Sharpe ratio of the momentum strategy. Risk-managed momentum is a much greater puzzle than the original version.

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% 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 because they are negatively correlated to those of the market and value factors. From 1927 to 2011, momentum had a monthly excess return of 1.75%, controlling for the Fama and French factors.[1] Moreover, momentum is not just a US stock market anomaly. Momentum has been shown 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

☆ We thank an anonymous referee, 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, Robert Kosowski, Soeren Hvidkjaer, and numerous seminar participants at BI Norwegian Business School, Copenhagen Business School, CUNEF (Colegio Universitario de Estudios Financieros), Imperial College London, Lubrafin (Luso-Brazilian Finance Meeting), Nova School of Business and Economics, National University of Singapore, Rotterdam School of Management, Universidad Carlos III de Madrid, University of Exeter, and University of New South Wales for their helpful comments and suggestions. A previous version of this paper circulated with the title "Managing the Risk of Momentum." Pedro Santa-Clara is supported by a grant from the Fundação para a Ciência e Tecnologia (PTDC/EGE-GES/101414/2008).

* Corresponding author.
E-mail addresses: p.barroso@unsw.edu.au (P. Barroso), psc@novasbe.pt (P. Santa-Clara).

[1] This result has led researchers to use momentum as an additional risk factor (Carhart, 1997).

[2] See Rouwenhorst (1998, 1999) for international evidence, Asness, Liew, and Stevens (1997) for country stock indices, Moskowitz and Grinblatt (1999) for industry portfolios, Okunev and White (2003) and Menkhof, Sarno, Schmeling, and Schrimpf (2012) for currency markets,

http://dx.doi.org/10.1016/j.jfineco.2014.11.010
0304-405X/

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 (WML) strategy delivered a −91.59% return in just two months.[4] In 2009, momentum experienced a crash of −73.42% 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.[5] They argue that hedging this time-varying market exposure produces stable momentum 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 autoregression of monthly realized variances yields an out-of-sample (OOS) R-square of 57.82%. This is 19.01 percentage points higher than a similar autoregression for the variance of the market portfolio, which is already famously predictable.[6]

Managing the risk of momentum leads to substantial economic gains. We simply scale the long-short portfolio by its realized volatility in the previous six months, targeting a strategy with constant volatility.[7] Scaling the portfolio to have constant volatility over time is a more natural way of implementing the strategy than having a constant amount in the long and short leg with varying volatility. This is widely accepted in industry, and targeting ex ante volatility is more common in practice than running constant leverage.[8] 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%; for risk-managed momentum, −28.40%. The maximum drawdown of raw momentum is −96.69% versus −45.20% for its risk-managed version.

The performance of scaled momentum is robust in subsamples and in international data. Managing the risk of momentum not only avoids its worse crashes but also improves the Sharpe ratio in the months without crashes. Risk management also improves the Sharpe ratio of momentum in all the major markets we examine: France, Germany, Japan, and the UK. When compared with plain momentum, risk management achieves a reduction in excess kurtosis and a less pronounced left skew in all of these markets.

Debate is ongoing about whether plain momentum per se is economically exploitable after transaction costs. For example, Lesmond, Schill, and Zhou (2004) infer costs indirectly from observed trading behavior and find that momentum is not exploitable. We do not address this debate directly, but we assess the impact of our risk management approach on transaction costs. Although the volatility scaling approach implies changes in leverage from month to month, we find that the turnover of the risk-managed strategy is very close to the turnover of the raw momentum strategy, so the transaction costs of both strategies are very similar. Given the much higher profitability of our strategy, transaction costs are less of a concern than for raw momentum.

One pertinent question is: Why does managing risk with realized variances work but using time-varying betas does not? To answer this question we decompose the volatility of momentum into a component related to the market (with time-varying betas) and a specific component. We find that the market component is only 23% of total risk on average, so most of the risk of momentum is specific to the strategy.[9] This specific risk is more persistent and predictable than the market component. The OOS R-square of predicting the specific component is 47.06% versus 20.87% for the market 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

(*footnote continued*)
Erb and Harvey (2006) for commodities, and Asness, Moskowitz, and Pederson (2013) for momentum across asset classes.

[3] See Daniel and Moskowitz (2012).

[4] Unless otherwise noted, by the performance of momentum we mean the return of the winners minus the return of the losers. The winners portfolio consists of those stocks in the top decile according to the distribution of cumulative returns from month $t-12$ to $t-2$. The losers portfolio is the group of stocks in the bottom decile of the same distribution. These returns can be found at Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french.

[5] Following negative returns for the overall market, winners tend to be low-beta stocks and the reverse for losers. Therefore, the winner-minus-losers strategy has a negative beta.

[6] See Engle and Bollerslev (1986) and Schwert (1989).

[7] This approach relies only on past data and thus does not suffer from look-ahead bias.

[8] We thank an anonymous referee for this insight.

[9] By 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.

Table 1
Descriptive statistics.
The long-run performance of momentum (WML, winners minus losers) is compared with the Fama and French risk factors: market (RMRF), size (SMB), and value (HML). All statistics are computed with monthly returns. Reported are the maximum and minimum one-month returns observed in the sample, the mean average excess return (annualized), the (annualized) standard deviation of each factor, excess kurtosis, skewness, and (annualized) Sharpe ratio. The sample returns are from 1927:03 to 2011:12.

Portfolio	Maximum	Minimum	Mean	Standard deviation	Kurtosis	Skewness	Sharpe ratio
RMRF	38.27	−29.04	7.33	18.96	7.35	0.17	0.39
SMB	39.04	−16.62	2.99	11.52	21.99	2.17	0.26
HML	35.48	−13.45	4.50	12.38	15.63	1.84	0.36
WML	26.18	−78.96	14.46	27.53	18.24	−2.47	0.53

their past residuals instead of gross returns produces a more stable version of momentum. Chaves (2012) shows that most of 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 explains the risk-managed momentum strategy. In Section 5, we decompose the risk of momentum and study the persistence of each of its components separately. In Section 6 we check if our findings also hold internationally. In Section 7, we assess the robustness of our findings across subsamples and briefly refer to other non-reported robustness results. Finally, Section 8 presents our conclusions.

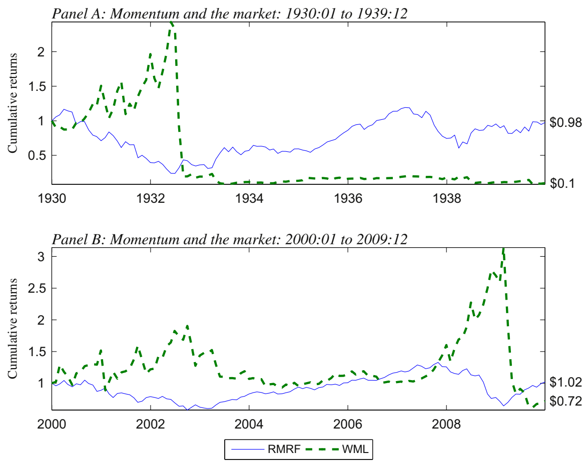

Fig. 1. Momentum crashes. The figure plots the cumulative return and terminal value of the momentum and market portfolio strategies in its two most turbulent periods: the 1930s and the 2000s. RMRF, market risk factor; WML, winners minus losers.

2. Momentum in the long run

We compare momentum with the Fama and French factors using a long sample of 85 years of monthly returns from July 1926 to December 2011 (see Appendix A for a description of the data). Daniel and Moskowitz (2012) use the same sample period.

Table 1 compares descriptive statistics for momentum in the long run with the Fama and French factors. Buying winners and shorting losers has provided large returns of 14.46% per year, with a Sharpe ratio higher than the market. The winners-minus-losers strategy offered an impressive performance for investors.

Nothing would be puzzling about momentum's returns if they corresponded to a very high exposure to risk. However, running an ordinary least squares (OLS) regression of the WML on the Fama and French factors gives (t-statistics in parentheses)

$$\begin{array}{lllll} r_{WML,t} = & 1.752 & -0.378 r_{RMRF,t} & -0.249 r_{SMB,t} & -0.677 r_{HML,t} \\ & (7.93) & (-8.72) & (-3.58) & (-10.76), \end{array} \quad (1)$$

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

The impressive excess returns of momentum, its high Sharpe ratio, and its 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, momentum has a dark side. 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, that is significant crash risk. Momentum returns can very rapidly turn into a free fall, wiping out decades of returns.

Fig. 1 shows the performance of momentum in the two most turbulent decades for the strategy: the 1930s and the 2000s. In July and August 1932, momentum had a cumulative return of −91.59%. From March to May 2009, momentum had another large crash of −73.42%. These short periods have an enduring impact on cumulative returns. For example, someone investing one dollar in the WML strategy in July 1932 would recover it only 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.[10] 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 world war II 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 (see, for example, Engle, 1982; Bollerslev, 1987). 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 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 $\{r_d\}_{d=1}^{D}$ be the daily returns and $\{d_t\}_{t=1}^{T}$ 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[11]

$$RV_{i,t} = \sum_{j=0}^{20} r^2_{i,d_t-j}. \tag{2}$$

Fig. 2 shows the monthly realized volatility of momentum. This varies substantially over time, from a minimum of 3.04% (annualized) to a maximum of 127.87%.

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

$$RV_{i,t} = \alpha + \rho RV_{i,t-1} + \varepsilon_t. \tag{3}$$

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 are available only from 1963:07 onward.

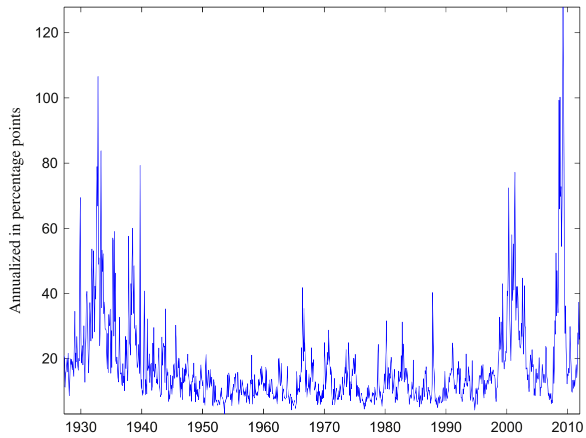

Fig. 2. The realized volatility of momentum obtained from daily returns in each month from 1927:03 to 2011:12.

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 out-of-sample (OOS) R-square uses the first 240 months to run an initial regression, so producing an OOS forecast. It then uses an expanding window of observations until 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 (market risk factor) and WML (winner minus losers) and add the same information for HML (high minus low) and SMB (small minus big). The last two columns show, respectively, the average realized volatility and its standard deviation.

Portfolio	α (t-statistic)	ρ (t-statistic)	R^2	OOS R^2	$\overline{\sigma}$	σ_σ
Panel A: 1927:03 to 2011:12						
RMRF	0.0010	0.60	36.03	38.81	14.34	9.97
	(6.86)	(23.92)				
WML	0.0012	0.70	49.10	57.82	17.29	13.64
	(5.21)	(31.31)				
Panel B: 1963:07 to 2011:12						
RMRF	0.0009	0.58	33.55	25.46	13.76	8.48
	(5.65)	(17.10)				
SMB	0.0004	0.33	10.68	−8.41	7.36	3.87
	(8.01)	(8.32)				
HML	0.0001	0.73	53.55	53.37	6.68	4.29
	(4.88)	(25.84)				
WML	0.0009	0.77	59.71	55.26	16.40	13.77
	(3.00)	(29.29)				

Momentum returns are the most volatile. From 1927:03 to 2011:12, the average realized volatility of momentum is 17.29, more than the 14.34 of the market portfolio. For 1963:07 onward, the average realized volatility of momentum is 16.40, also the highest when compared with RMRF, SMB, and HML.

In the full sample period, the standard deviation of monthly realized volatilities is higher for momentum (13.64) than the market (9.97). Panel B confirms this result compared with the other factors in the 1963:07 onward 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

[10] Daniel and Moskowitz (2012) argue this is due to the option-like payoffs of distressed firms in bear markets.

[11] Correcting for serial correlation of daily returns does not change the results significantly.

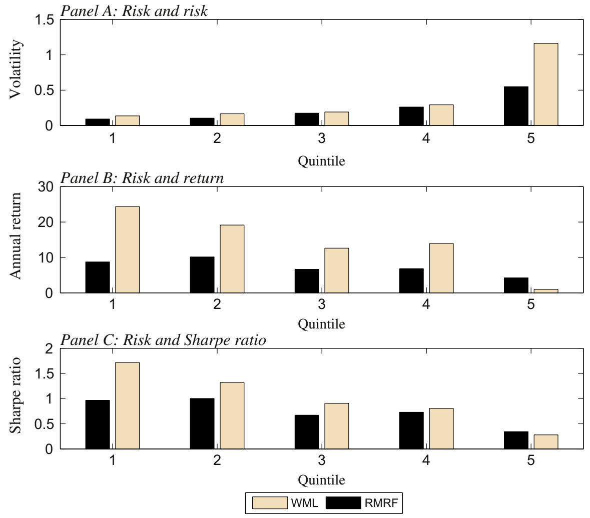

Fig. 3. The performance of the market factor (RMRF) and momentum (WML) conditional on realized volatility in the previous six months. Returns are sorted into quintiles according to volatility in the previous six months for each factor. The figure presents the following year volatility (computed from daily returns), the cumulative return in percentage points, and the Sharpe ratio. The returns are from 1927:03 to 2011:12.

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 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 with the accuracy of the historical mean $\overline{RV}_{i,t}$. As a measure of goodness of fit, we estimate the OOS R-square as

$$R^2_{i,OOS} = 1 - \frac{\sum_{t=S}^{T-1}(\widehat{\alpha}_t + \widehat{\rho}_t RV_{i,t} - RV_{i,t+1})^2}{\sum_{t=S}^{T-1}(\overline{RV}_{i,t} - RV_{i,t+1})^2}, \tag{4}$$

where S is the initial training sample and $\widehat{\alpha}_t, \widehat{\rho}_t$, and $\overline{RV}_{i,t}$ are estimated with information available only up to time t.

The last column of Table 2 shows the OOS R-squares of each autoregression. The AR(1) of the realized variance of momentum has an OOS R-square of 57.82% (full sample), which is 50% 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.

Fig. 3 illustrates the potential of realized variance of momentum to condition exposure to the factor. We sort the months into quintiles according to the level of realized variance in the previous six 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, no obvious trade-off exists between risk and return. This illustrates the well-known difficulty in estimating this relation (see, for example, Glosten, Jagannathan, and Runkle, 1993; Ghysels, Santa-Clara, and Valkanov, 2005). For momentum the data show a negative relation between risk and return.[12]

As a result, 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 Section 4, 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.

4. Risk-managed momentum

We use an estimate of momentum risk to scale the exposure to the strategy to have constant risk over time.[13] For each month we compute a variance forecast $\hat{\sigma}_t^2$ from daily returns in the previous six months.[14] Let $\{r_{WML,t}\}_{t=1}^{T}$ be the monthly returns of momentum and $\{r_{WML,d}\}_{d=1}^{D}, \{d_t\}_{t=1}^{T}$ be the daily returns and the time series of the dates of the last trading sessions of each month.

The variance forecast is

$$\hat{\sigma}^2_{WML,t} = 21 \sum_{j=0}^{125} r^2_{WML,d_{t-1}-j}/126. \tag{5}$$

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:

$$r_{WML^*,t} = \frac{\sigma_{\text{target}}}{\hat{\sigma}_t} r_{WML,t}, \tag{6}$$

where $r_{WML,t}$ is the unscaled or plain momentum, $r_{WML^*,t}$ is the scaled or risk-managed momentum, and σ_{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 the strategy is still self-financing. We pick a target corresponding to an annualized volatility of 12%.[15]

Fig. 4 shows the weights of the scaled momentum strategy over time, interpreted as the dollar amount in the

[12] Wang and Xu (2011) and Tang and Mu (2012) find a similar result forecasting the return of momentum with the volatility of the market.

[13] Volatility-scaling has been used in the time series momentum literature (Moskowitz, Ooi, and Pederson, 2012; 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 show that the persistence in risk of the winners-minus-losers strategy is more interesting than that of a long-only portfolio of assets, such as the market.

[14] We also used one-month and three-month realized variances as well as exponentially weighted moving average (EWMA) with half-lifes of one, three, and six months. All work well with nearly identical results.

[15] The annualized standard deviation from monthly returns is higher than 12% as volatilities at daily frequency are not directly comparable with those at lower frequencies due to the small positive autocorrelation of daily returns.

Table 3
Momentum and the economic gains from scaling.
The first row presents as a benchmark the economic performance of plain momentum (*WML*) from 1927:03 to 2011:12. The second row presents the performance of risk-managed momentum (*WML**). The risk-managed momentum uses the realized variance in the previous six months to scale the exposure to momentum. The mean, the standard deviation, the Sharpe ratio, and the information ratio are annualized. To obtain an information ratio that does not depend on the volatility target we divided previously both (*WML*) and (*WML**) by their respective standard deviations.

Portfolio	Maximum	Minimum	Mean	Standard deviation	Kurtosis	Skewness	Sharpe ratio	Information ratio
WML	26.18	−78.96	14.46	27.53	18.24	−2.47	0.53	–
*WML**	21.95	−28.40	16.50	16.95	2.68	−0.42	0.97	0.78

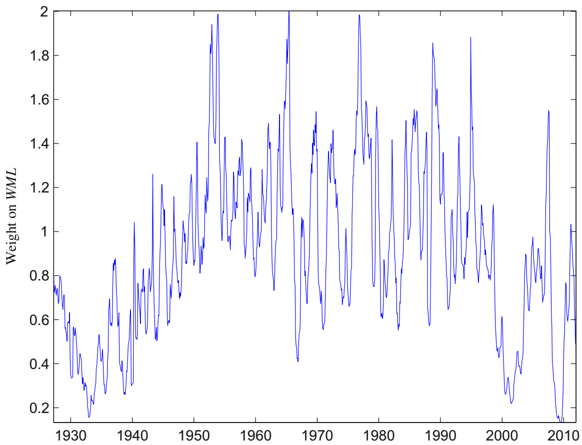

Fig. 4. Weights of the scaled momentum, 1927:03–2011:12. The risk-managed momentum uses the realized variance in the previous six months to scale the exposure to momentum (*WML*).

long or short leg. These range between the values of 0.13 and 2.00, reaching the most significant lows in the early 1930s, in 2000–02, and in 2008–09. On average, the weight is 0.90, slightly less than full exposure to momentum.

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 information ratio of *WML** compared with conventional momentum has a very high value of 0.78.

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. Fig. 5 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.

Also, the risk-managed strategy no longer has variable and persistent risk, so risk management does work.[16]

The benefits of risk management are especially important in turbulent times. Fig. 6 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 1930s. This compares very favorably with the pure momentum strategy which loses 90% in the same period. In the 2000s simple momentum loses 28% of wealth, because of the crash in 2009. Risk-managed momentum ends the decade up 88% as it not only avoids the crash but also captures part of the positive returns of 2007–2008.

Risk-managed momentum depends only on ex ante information, so this strategy could be implemented in real time. Running a long-short strategy to have constant volatility is closer to what real investors (such as hedge funds) try to do than keeping a constant amount invested in the long and short legs of the strategy.

One relevant issue is whether time-varying weights induce such an increase in turnover that eventually offsets the benefits of the strategy after transaction costs. To control for this, we compute the turnover of momentum and its risk-managed version from stock-level data on returns and firm size from 1951:03 to 2010:12.[17] We find that the turnover of momentum is 74% per month and the turnover of risk-managed momentum is 75%. The increase in turnover is low because $\hat{\sigma}_t$, with an AR(1) coefficient of 0.97, is highly persistent from month to month. This increase in turnover is not sufficient to offset the benefits of volatility scaling. As in Grundy and Martin (2001), we calculate the round-trip cutoff cost that would render the profits of each strategy insignificant at the 5% level. We find that cost to be 1.27 percentage points for *WML* and 1.75 percentage points for *WML**. So the transaction costs that would remove the significance of the profits of risk-managed momentum are 38% higher than for conventional momentum.

5. Anatomy of momentum risk

A well-documented result in the momentum literature is that momentum has time-varying market betas (Grundy and Martin, 2001). This is an intuitive finding because, 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

[16] 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 autocorrelation of raw momentum is significant up to 15 lags but it is only 1 lag for risk-managed momentum. So persistence in risk is much smaller for the risk-managed strategy.

[17] We explain the computation in Appendix B. The different period considered relative to the rest of the analysis is due to data availability.

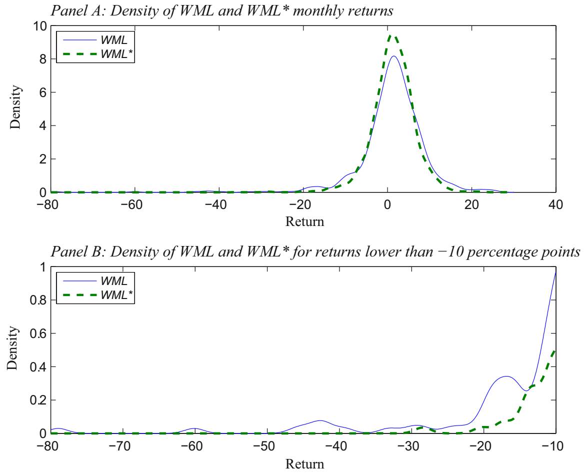

Fig. 5. The density of plain momentum (*WML*) and risk-managed momentum (*WML**). The risk-managed momentum uses the realized variance in the previous six months to scale the exposure to momentum. The returns are from 1927:03 to 2011:12.

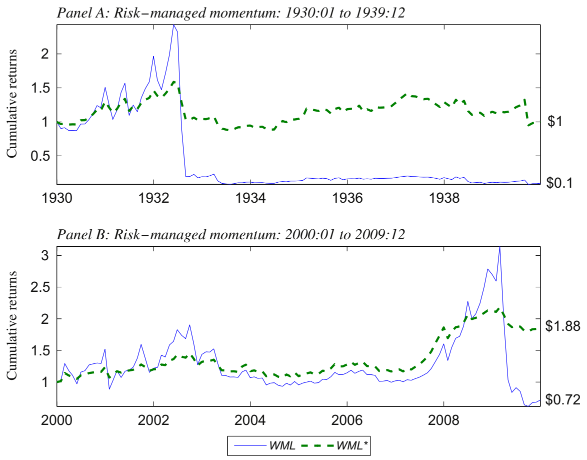

Fig. 6. The benefits of risk-management in the 1930s and the 2000s. The risk-managed momentum (*WML**) uses the realized variance in the previous six months to scale the exposure to momentum (*WML*).

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:

$$RV_{wml,t} = \beta_t^2 RV_{rmrf,t} + \sigma_{e,t}^2. \quad (7)$$

The realized variances and betas are estimated with six months of daily returns. On average, the market component $\beta_t^2 RV_{rmrf,t}$ accounts for only 23% of the total risk of momentum. Almost 80% of the momentum risk is specific to the strategy. Also, the different components do not have the same degree of predictability. Table 4 shows the results of an AR(1) on each component of risk.

Table 4

Decomposition of the risk of momentum.

Each row shows the results of an AR(1) for six-month, non-overlapping periods. The first row is for the realized variance of the WML (winners minus losers); 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 (market risk factor). The fourth row is the systematic component of momentum risk; the last row, the specific component. The out of sample *R*-squares use an expanding window of observations after an initial in-sample period of 20 years.

Variable	α (t-statistic)	ρ (t-statistic)	R^2	R^2_{OOS}
σ^2_{wml}	0.0012	0.70	48.67	43.82
	(2.59)	(12.58)		
σ^2_{rmrf}	0.0012	0.50	24.53	6.70
	(4.29)	(7.37)		
β^2	0.3544	0.21	4.59	5.33
	(6.05)	(2.83)		
$\beta^2\sigma^2_{rmrf}$	0.0007	0.47	21.67	20.87
	(2.73)	(6.80)		
σ^2_{ε}	0.0007	0.72	52.21	47.06
	(2.69)	(13.51)		

Either in-sample or out-of-sample, β_t^2 is the least predictable component of momentum risk. Its OOS *R*-square is only 5.33%. The realized variance of the market also has a small OOS *R*-square of 6.70%. When combined, both elements form the market risk component and show more predictability with an OOS *R*-square of 20.87%, but still less than the realized variance of momentum with an OOS *R*-square of 43.82%. The most predictable component of momentum variance is the specific risk with an OOS *R*-square of 47.06%, 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.[18]

6. 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 most 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

[18] 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.

Table 5
The international evidence.

The performance of plain momentum (*WML*) and scaled momentum (*WML**) in the major non-US markets. Original data are from Datastream. The returns for country-specific momentum portfolios are from Chaves (2012). The scaled momentum uses the realized volatility in the previous six months, obtained from daily data. The information ratio of the scaled momentum takes plain momentum as the benchmark. The mean return, the standard deviation, the Sharpe ratio, and the information ratio are all annualized. To obtain an information ratio that does not depend on the volatility target we divided previously both *WML* and *WML** by their respective standard deviations. The returns are from 1980:07 to 2011:10.

Statistic	France		Germany		Japan		UK	
	WML	*WML**	*WML*	*WML**	*WML*	*WML**	*WML*	*WML**
Maximum	35.43	31.88	22.04	19.34	17.21	16.61	23.23	45.98
Minimum	−28.11	−15.50	−22.56	−11.11	−29.51	−21.56	−36.47	−22.23
Mean	13.19	17.11	18.37	21.02	1.65	4.21	18.86	40.46
Standard deviation	19.66	16.57	18.10	15.08	19.64	17.88	17.24	22.81
Kurtosis	6.85	5.17	4.02	1.45	4.01	1.96	11.78	4.95
Skewnewss	−0.16	0.73	−0.05	0.33	−0.83	−0.62	−1.43	0.56
Sharpe ratio	0.67	1.03	1.02	1.39	0.08	0.24	1.09	1.77
Information ratio	–	0.90	–	0.73	–	0.34	–	1.19

each country. These are defined as the difference between the return of the winner quintile and the loser quintile.[19] Only four countries consistently satisfy the minimum data requirements: France, Germany, Japan, and the UK. We choose to report results only for these four countries for which data are more reliable and, therefore, avoid possible issues with missing observations and outliers.[20]

Table 5 shows the results of the risk-managed momentum in the four countries considered. The risk-managed strategy uses the realized volatility in the previous six months to target a constant volatility of 12% (annualized).

Managing the risk of momentum improves the mean return in all markets while reducing the standard deviation in three of the four markets (France, Germany, and Japan). The Sharpe ratio increases in all markets. 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 as much as 0.68 in the UK.

The benefits of risk management are especially strong in reducing higher-order risk. The excess kurtosis is smaller in all four markets after managing risk, and the skewness of momentum becomes less negative or even turns positive. Raw momentum has negatively skewed returns in all markets, and risk-managed momentum has a positive skew in France, Germany, and the UK. The information ratio of risk-managed momentum ranges between 0.34 for Japan and 1.19 in the UK. So risk management adds considerable value when compared with the plain momentum strategy in all countries.

As such, we conclude that the benefits of managing the risk of momentum are pervasive in international data. Because we did not have these 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.

7. Robustness checks

As Fig. 1 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. Our sample has two crashes of a very large magnitude: the first one in 1932 and the second in 2009. Therefore, it is pertinent to ask to what extent are our entire results driven by these two singular occurrences (Table 6).

To address this issue, we examine the performance of both momentum and risk-managed momentum in different subsamples. First, in a simple robustness exercise, we split the sample in two halves. This shows the results are not driven entirely by just one of the crashes. But as the two halves include a major crash, we also examine a sample of the entire span from 1927:03 to 2011:12 but excluding the years of these two crashes. Finally, we consider the relatively benign period of 1945:01 to 2005:12. This roughly corresponds to the post-war period up to the years preceding the Great Recession, a period in which momentum performed very well.

In all samples, risk management reduces excess kurtosis and left skewness. Excess kurtosis drops from 23.57 to 4.53 in the first half of the sample and from 7.15 to 2.47 in the no-crash sample. Skewness increases from −0.91 to −0.17 in the post-war sample and from −3.05 to −0.72 in the first half of the sample. So risk management reduces higher-order risk, even in samples without a major crash.

The increase in Sharpe ratio is between 0.29 in the post-war sample and 0.47 in the second half of the sample. Even in the samples without a major crash, risk-managed momentum shows a very high information ratio of at least 0.60.[21]

[19] We thank Denis Chaves for letting us use his data. See his paper for a more detailed description of the international data used.

[20] 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.

[21] We thank an anonymous referee for this suggestion.

Table 6
Performance of plain momentum (*WML*) and scaled momentum (*WML**) in different subsamples

The first half of the sample is from 1927:03 to 1969:12. The second is from 1970:01 to 2011:12. The no-crash sample is from 1927:03 to 2011:12, excluding the years of 1932 and 2009. The post-war sample is from 1945:01 to 2005:12. The Sharpe ratio, standard deviation, information ratio, and the mean returns are all annualized. To obtain an information ratio that does not depend on the volatility target we divided previously both *WML* and *WML** by their respective standard deviations.

Statistic	First half		Second half		No-crash		Post-war	
	WML	*WML**	*WML*	*WML**	*WML*	*WML**	*WML*	*WML**
Maximum	24.89	19.76	26.18	21.95	26.18	21.95	26.18	21.95
Minimum	−78.96	−28.40	−45.89	−17.00	−43.94	−28.40	−42.09	−17.00
Mean	13.12	13.96	15.83	19.08	17.34	17.39	17.30	19.61
Standard deviation	29.33	16.29	25.59	17.58	23.61	16.74	20.02	16.97
Kurtosis	23.57	4.53	7.62	1.19	7.15	2.47	6.01	1.20
Skewnewss	−3.05	−0.72	−1.51	−0.19	−1.13	−0.33	−0.91	−0.17
Sharpe ratio	0.45	0.86	0.62	1.09	0.73	1.04	0.86	1.16
Information ratio	–	0.73	–	0.81	–	0.60	–	0.64

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

We show that the risk of momentum is highly variable and predictable. Another relevant question is whether this predictability is specific to the predictive variable used (momentum's own lagged risk). In unreported results, we find that forecasting the risk of momentum with the realized variance of the market in the previous six months produces very similar results. So there are alternative methods to manage the risk of momentum and exploit its predictability.[22]

Another issue is to what extend our results overlap with other research on the predictability of momentum's risk. Most of that literature has focused on the time-varying beta 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 risk that is much greater and more robust than the bear-market indicator. (We omit these results for the sake of brevity.)

8. Conclusion

Unconditional momentum has a distribution that is far from normal, with huge crash risk. 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 subsamples. The transaction costs needed to remove the significance of risk-managed momentum profits are nearly 40% higher than for conventional momentum.

[22] We thank an anonymous referee for pointing this out.

Appendix A. 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 are from July 1926 to December 2011, and the daily data are 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 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, each bin has an equal number of NYSE firms. The WML strategy is to short the lowest (loser) decile and take 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 and French momentum portfolios for the 1963:07–2011:12 sample.

Appendix B. Turnover

The turnover of a leg of the momentum portfolio is

$$x_t = 0.5 \times \sum_{i}^{N_t} \left| w_{i,t} - \tilde{w}_{i,t-1} \right|, \tag{8}$$

where $w_{i,t}$ is the weight of stock i in the leg of the portfolio at time t, N_t is the number of stocks in the leg of the portfolio at time t, $r_{i,t}$ is the return of asset i at time t, and $\tilde{w}_{i,t-1}$ is the weight in the current period right before trading

$$\tilde{w}_{i,t-1} = \frac{w_{i,t-1}(1+r_{i,t})}{\sum_i^{N_t} w_{i,t-1}(1+r_{i,t})}. \tag{9}$$

The turnover of the *WML* is simply the sum of the turnover of the short and the long leg. In the case of the risk-managed portfolio, we compute for each leg the turnover as

$$x_t = 0.5 \times \sum_i^{N_t} \left| \frac{w_{i,t}}{L_t} - \frac{\tilde{w}_{i,t-1}}{L_{t-1}} \right|, \tag{10}$$

where $L_t = \sigma_{\text{target}}/\hat{\sigma}_t$. The turnover of WML^* is simply the sum of the turnover of the long with the short leg.

References

Asness, C., 2011. Momentum in Japan: the exception that proves the rule. Journal of Portfolio Management 37, 67–75.

Asness, C., Liew, J., Stevens, R., 1997. Parallels between the cross-sectional predictability of stock and country returns. Journal of Portfolio Management 23, 79–87.

Asness, C., Moskowitz, T., Pedersen, L., 2013. Value and momentum everywhere. Journal of Finance 58, 929–986.

Baltas, A., Kosowski, R., 2013. Momentum strategies in futures markets and trend-following funds. European Financial Data Institute – French Finance Association (EUROFIDAI-AFFI) paper. Finance meeting, Paris, France, December 2012.

Barroso, P., 2012. The bottom-up beta of momentum. Unpublished working paper. University of New South Wales, Sydney.

Blitz, D., Huij, J., Martens, M., 2011. Residual momentum. Journal of Empirical Finance 18, 506–521.

Bollerslev, T., 1987. A conditional heteroskedastic time series model for speculative prices and rates of return. Review of Economics and Statistics 69, 542–547.

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

Chaves, D., 2012. Eureka! A momentum strategy that also works in Japan. Unpublished working paper. Research Affiliates, Newport Beach, California.

Cooper, M., Gutierrez, R., Hameed, A., 2004. Market states and momentum. Journal of Finance 59, 1345–1366.

Daniel, K., Moskowitz, T., 2012. Momentum crashes. Unpublished working paper. Columbia University, New York.

Engle, R., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. Econometrica 50, 987–1008.

Engle, R., Bollerslev, T., 1986. Modeling the persistence of conditional variances. Econometric Reviews 5, 1–50.

Erb, C., Harvey, C., 2006. The strategic and tactical value of commodity futures. Financial Analysts Journal 62, 69–97.

Fama, E., French, K., 1992. The cross section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E., French, K., 2012. Size, value, and momentum in international stock returns. Journal of Financial Economics 105, 457–472.

Ghysels, E., Santa-Clara, P., Valkanov, R., 2005. There is a risk-return trade-off after all. Journal of Financial Economics 76, 509–548.

Glosten, L., Jagannathan, R., Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48, 1779–1801.

Griffin, J., Ji, X., Martin, J., 2003. Momentum investing and business cycle risk: evidence from pole to pole. Journal of Finance 53, 2515–2546.

Grinblatt, M., Titman, S., 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 394–415.

Grinblatt, M., Titman, S., 1993. Performance measurement without benchmarks: an examination of mutual fund returns. Journal of Business 66, 47–68.

Grundy, B., Martin, J., 2001. Understanding the nature of the risks and the source of the rewards to momentum investing. Review of Financial Studies 14, 29–78.

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

Lesmond, D., Schill, M., Zhou, C., 2004. The illusory nature of momentum profits. Journal of Financial Economics 71, 349–380.

Menkhoff, L., Sarno, L., Schmeling, M., Schrimpf, A., 2012. Currency momentum strategies. Journal of Financial Economics 106, 620–684.

Moskowitz, T., Ooi, Y., Pedersen, L., 2012. Time series momentum. Journal of Financial Economics 104, 228–250.

Moskowitz, T., Grinblatt, M., 1999. Do industries explain momentum? Journal of Finance 54, 1249–1290.

Okunev, J., White, D., 2003. Do momentum-based strategies still work in foreign currency markets? Journal of Financial and Quantitative Analysis 38, 425–447.

Rouwenhorst, K., 1998. International momentum strategies. Journal of Finance 53, 267–284.

Rouwenhorst, K., 1999. Local return factors and turnover in emerging stock markets. Journal of Finance 54, 1439–1464.

Schwert, G., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115–1154.

Tang, F., Mu, J., 2012. Market volatility and momentum. Unpublished working paper. University of Texas, Austin.

Wang, K., Xu, W., 2011. Market volatility and momentum. Unpublished working paper. University of Toronto, Toronto.