Momentum Profits
and Non-Normality Risks

January 2007

Ana-Maria Fuertes
Associate Professor of Finance, Cass Business School, City University London

Joëlle Miffre
Professor of Finance, EDHEC Business School

Wooi-Hou Tan
Director, Cyberring Ltd, London
ABSTRACT
The paper examines the role of non-normality risks in explaining the momentum puzzle of equity returns. It shows that momentum returns are not normally distributed. About 70 basis points of the annual momentum profits can be attributed to systematic skewness risk. This finding is pervasive across nine strategies and is reinforced when time dependencies in abnormal returns and risks are explicitly modeled. The analysis also reveals that the market and skewness risks of momentum portfolios evolve over the business cycle in a manner that is consistent with market timing and risk aversion. While the momentum puzzle still remains, these findings are in line with market efficiency.

Keywords: Momentum strategy; Abnormal returns; Skewness; Conditional asset pricing.
JEL classifications: G12, G14.

Authors’ Note: At the time of writing, Joëlle Miffre was Associate Professor of Finance at Cass Business School.

Acknowledgements: We appreciate the comments of Chris Brooks, Jerry Coakley, Keith Cuthbertson, Soosung Hwang, Aneel Keswani and seminar participants at Cass Business School and the Institute for Advanced Studies. We would also like to thank Cass Business School for its financial support.

EDHEC is one of the top five business schools in France owing to the high quality of its academic staff (104 permanent lecturers from France and abroad) and its privileged relationship with professionals that the school has been developing since its establishment in 1906. EDHEC Business School has decided to draw on its extensive knowledge of the professional environment and has therefore concentrated its research on themes that satisfy the needs of professionals.

EDHEC pursues an active research policy in the field of finance. Its “Risk and Asset Management Research Centre” carries out numerous research programs in the areas of asset allocation and risk management in both the traditional and alternative investment universes.

Copyright © 2007 Edhec
1. INTRODUCTION

The profitability of momentum or relative-strength strategies has become a stylized fact in finance. Zero-cost portfolios that are long recent winners and short recent losers generate positive returns over holding periods that range from 3 to 12 months. Price continuation has been identified over time (Jegadeesh and Titman, 1993, 2001), across countries (Rouwenhorst, 1998; Chui et al., 2000; Griffin et al., 2003) and across industries (Moskowitz and Grinblatt, 1999). Recent studies suggest that momentum strategies also perform well across equity styles (Chen and De Bondt, 2004) and across asset classes (Okunev and White, 2003; Miffre and Rallis, 2005). Against this wealth of evidence, momentum profits can hardly be attributed to data mining.

Market-efficiency advocates argue that momentum returns are a rational compensation for risk, a liquidity premium and/or an illusion induced by market frictions. While momentum profits may well be a misconception triggered by transaction costs (Lesmond et al., 2004) or a premium for illiquidity (Sadka, 2003), the evidence in favour of the risk-based rationale is at best weak. Momentum profits indeed persist after accounting for market risk (Jegadeesh and Titman, 1993), size and book-to-market value (Fama and French, 1996), and macroeconomic and financial factors (Griffin et al., 2003). Likewise, the returns of momentum strategies do not seem to be a compensation for cross-sectional variation in unconditional expected returns (Conrad and Kaul, 1998) as they tend to mean revert beyond one year (Jegadeesh and Titman, 2001). There is also ample evidence to suggest that momentum profits are not merely a reward for time-varying risks (Wu, 2002; Griffin et al., 2003; Karolyi and Kho, 2004). Using bootstrap simulations, Karolyi and Kho (2004) test whether a range of return-generating models proposed in the literature can explain the profitability of momentum strategies. Overall, none of the models is able to generate momentum returns as large as the actual profits over the sample period.

Behavioural economists take a different view. They interpret momentum as a consequence of the market responding slowly to news. They explain the under- and over-reaction observed in equity markets in terms of psychology (cognitive biases) and limits to arbitrage. Barberis et al. (1998) show that conservatism leads to momentum while representativeness heuristic leads to overvaluation followed by price correction. Daniel et al. (1998) relate the predictability of equity returns to biased self-attribution and overconfidence. Hong and Stein (1999) attribute momentum and reversal partly to bounded rationality. Evidence that momentum portfolios generate negative returns over holding periods beyond one year (Jegadeesh and Titman, 2001) lends support to the behavioural hypothesis: transactions by short-term momentum traders temporarily move prices away from long-term equilibrium, causing them to overreact. Once the overreaction is acknowledged by the market, a correction sets in. Sophisticated investors who recognize these behavioural biases can consistently earn higher returns without taking more risk.

This paper re-addresses the role of systematic risks in explaining the momentum puzzle. The three-factor model of Fama and French (1993) has become a widely used tool in security analysis. Yet the two systematic risk factors over and above market risk proposed by Fama and French (1993), capitalization (firm size) and book-to-market, are motivated on an empirical basis. It is well known that size and book-to-market do not stand up to theoretical scrutiny. In sharp contrast, it is widely recognized that investors have an aversion to risk as measured by the standard deviation of returns and a preference for positive skewness and negative excess kurtosis. Besides, it is commonly acknowledged that the first two moments fail to fully describe the distribution of portfolio returns. For example, departures from normality have been evidenced for emerging market indices (Harvey, 1995), hedge fund indices (Agarwal and Naik, 2004), relative-strength strategies (Harvey and Siddique, 2000), and futures contracts (Christie-David and Chaudhry, 2001). Finally, and perhaps most importantly, systematic skewness (also called coskewness), as defined by the contribution of an asset to the skewness of a well-diversified portfolio, is relevant to asset valuation (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000; Chung et al., 2004).

Against this background, it seems only natural to investigate the extent to which the profitability of relative-strength strategies is a compensation for exposure to systematic departures from normality.

1 - Rouwenhorst (1998) extends the U.S. evidence in Jegadeesh and Titman (1993, 2001) to European markets while Chui et al. (2000) detect price continuation in Asian markets with Japan as an exception. For a sample of 40 countries, Griffin et al. (2003) show that the evidence from developed markets is stronger than that of emerging markets. Moskowitz and Grinblatt (1999) argue that industry momentum encompasses stock momentum. Against this background, it seems only natural to investigate the extent to which the profitability of relative-strength strategies is a compensation for exposure to systematic departures from normality.

2 - Chen and De Bondt (2004) document that styles that were in favor (out of favor) over the recent past, perform better (worse) in the near future. Okunev and White (2003) argue that momentum strategies perform well in foreign exchange markets, whereas Miffre and Rallis (2005) detect a profitable momentum pattern in commodity futures markets.

3 - Lesmond et al. (2004) show that trading costs can wipe out the momentum profits. The momentum strategies are indeed highly trading intensive and pick relatively illiquid stocks with large trading costs. Moreover, the profits are mainly driven by the losers that need to be sold short (Hong et al., 2000).

4 - Harvey and Siddique (2006) highlight agency problems (Brennan, 1993) and limited liability (Golec and Tamarkin, 1998) as factors that could induce skewness in ex post equity returns.
This paper contributes to the literature in three directions. First, it explores through a time-series framework the hypothesis that the momentum profits are, at least in part, due to non-normality in the return distribution of relative-strength strategies. It complements the work by Harvey and Siddique (2000) which instead focuses on explaining the cross-sectional variation of expected returns. We confirm their finding that winner returns are more negatively skewed than loser returns. In addition, we document that the winners exhibit higher kurtosis than the losers. In principle, this could suggest that the market compensates investors with higher returns for exposure to the relatively low (negative) skewness and high kurtosis of momentum returns, a rationale consistent with risk aversion and rational pricing in an efficient market. To investigate this conjecture, we augment the single index model and the Fama and French (1993) model with the returns of skewness- and kurtosis-mimicking portfolios. The evidence suggests that the momentum effect is partly driven by undesired negative coskewness with the market portfolio. Most importantly, we show that the momentum profitability falls when systematic skewness risk is controlled for. These findings are shown to be pervasive across nine momentum-based trading strategies.

Second, the paper explicitly deals with several specification problems in the asset pricing models to provide more robust evidence on the importance of non-normality risks. Time-varying idiosyncratic risk is explicitly captured through a GARCH(1,1) process as in Karolyi and Kho (2004). Another model specification, following Wu (2002), is used to acknowledge the fact that the dynamic asset allocation of momentum traders induces time dependence in the systematic risks and in the abnormal performance of relative-strength strategies. Such time variation may also be linked to the inexorable succession of bull and bear markets. The relevance of these model specifications is borne out by tests that uncover not only volatility clustering in the asset-pricing model residuals but also time dependence in market, book/market value and skewness risks. The main finding on the importance of systematic skewness risk is shown to remain unchallenged when such time dependencies are controlled for. Besides, once skewness risk is accounted for in the asset-pricing models, the momentum profitability decreases significantly. On average, skewness risk is found to contribute 70 basis points towards the annualized abnormal momentum returns.

Third, our analysis uncovers that two of the systematic risks considered, market and skewness risks, evolve over the business cycle in a manner that is consistent with risk aversion and market timing. In particular, the risks of the winner (loser) portfolios appear to be higher (lower) during economic expansions than during economic recessions. This suggests that the asset allocation of a momentum trader can be cast as that of a market timer who tactically tilts his asset allocation towards the best performing asset class.

The evidence that part of the momentum profits is due to systematic skewness risk is not only theoretically appealing to market-efficiency advocates, since it explains part of the momentum puzzle, but can also be given an intuitive interpretation. Suppose that the winner is a company that has disclosed a series of exceptionally high earnings in the recent past. While it is expected to sustain such level of earnings in the near future, there is a small but non-negligible probability that the company fails to deliver. Indeed, as high earnings cannot be sustained forever, price continuation is eventually followed by mean reversion (Jegadeesh and Titman, 2001).

The remainder of the paper is organized as follows. Section 2 describes the dataset and outlines the methodology adopted to construct momentum portfolios and systematic skewness- and kurtosis-mimicking portfolios. Section 3 discusses summary statistics on the winner, loser and momentum portfolio returns. Section 4 reports the empirical findings from non-normality augmented versions of the single index model and the Fama-French model. It highlights that some of the momentum profits are a compensation for exposure to undesired systematic skewness. Section 5 re-addresses the momentum puzzle in the context of asset-pricing models that allow for time variation in the idiosyncratic risk, systematic risks and abnormal performance. Section 6 concludes.
2. DATA AND METHODOLOGY
This section discusses the data and the methodology that we adopt to construct momentum portfolios and mimicking portfolios for systematic skewness and kurtosis. It also outlines the risk-adjusted profitability measures considered and the hypotheses to be tested.

2.1. Momentum portfolios
Monthly stock prices over the period 31/01/1973 to 31/08/2004 are obtained from Datastream for all the stocks listed on the Amex, NYSE and NASDAQ exchanges. To mitigate problems of survivorship bias we also include delisted stocks. The cross-section ranges from a low of 2535 stocks in January 1973 to a peak of 9456 in June 2001.

We consider the 9 momentum strategies that arise from combining ranking (R) and holding (H) periods of 3, 6 and 12 months. At the end of each month, stocks are sorted into deciles based on their return over the previous R months. The performance of both the top- and bottom-decile portfolios is tracked over the subsequent H months (the stocks in each decile are equally weighted). This trading rule is called the R-H momentum strategy.

We follow Moskowitz and Grinblatt (1999) and Jegadeesh and Titman (2001) in forming overlapping winner (and loser) portfolios. More specifically, taking the 6-3 strategy as example, the return on the winner portfolio in, say, April equals the equally-weighted return in April of the 3 top deciles portfolios that were formed at the end of January, February, and March. The same approach is used for the loser portfolio. We follow Jegadeesh and Titman (2001) in filtering out stocks priced below $5. The return on the self-financing momentum strategy is defined as the difference between winner and loser returns. This approach is rolled forward one month at a time to form a sequence of winner/loser portfolio returns.

2.2. Systematic skewness- and kurtosis-mimicking portfolios
In order to test whether momentum returns are a reward for exposure to systematic skewness and kurtosis risks, we augment standard asset-pricing models with the returns of mimicking portfolios for skewness and kurtosis. The methodology that we employ to construct the latter is a direct extension of that proposed in Fama and French (1993) to form size and book/market value sorted portfolios. Let $r_{it}$ denote the return on stock $i$ at month $t$ and $r_{Mt}$ the return on the S&P500 Composite index. We take the view that, in a well-diversified portfolio, idiosyncratic skewness (kurtosis) is eliminated and thus that investors only earn compensation for exposure to systematic skewness (kurtosis). We calculate the systematic skewness or coskewness (with the market portfolio) of each stock as:

$$\text{coSK}_i = \frac{E[(r_{it} - \bar{r}_i)(r_{Mt} - \bar{r}_M)^2]}{E[(r_{Mt} - \bar{r}_M)^2]} \quad (1)$$

where the unconditional means, $\bar{r}_i$ and $\bar{r}_M$, and the expectations are measured over the 60-month window $[t-59,...,t]$.

We then sort the stocks according to $\text{coSK}_i$ and form two portfolios that contain the 30% of stocks with the lowest and highest $\text{coSK}_i$ respectively. The difference in the average return on these low and high coskewness portfolios provides the return on the skewness-mimicking portfolio in each of the subsequent $[t+1,...,t+12]$ months. The 60-month window to calculate (1) is rolled forward 12-months to form new portfolios. This recursive approach yields a monthly time series of 320 skewness-mimicking portfolio returns.

The kurtosis-mimicking portfolios are formed using a similar approach with two differences. First, stocks are sorted according to their systematic kurtosis or cokurtosis

$$\text{coKU}_i = \frac{E[(r_{it} - \bar{r}_i)(r_{Mt} - \bar{r}_M)^4]}{E[(r_{Mt} - \bar{r}_M)^4]} \quad (2)$$

which is the counterpart of (1) for the fourth moment. Second, the kurtosis-mimicking portfolio return is the difference in the return on the high- and low-kurtosis portfolios.

6. The effective sample starts in January 1978 because the month-t returns of the skewness- and kurtosis-mimicking portfolios are based on stock rankings over $[t-60, t-1]$.

7. The return is in say April is only one third determined by the winner/loser portfolios formed in March and therefore the momentum returns are unlikely to be driven by bid-ask bounce. Hence, as in Moskowitz and Grinblatt (1999) we do not skip one month between the ranking and holding periods.

8. Note that the skewness- and kurtosis-mimicking portfolios and the momentum portfolios are based on the same data set. The only difference is that, as in Fama and French (1993), the mimicking portfolios are constructed without setting a minimum-price filter.

9. We build on the unconditional three-moment CAPM developed by Kraus and Litzenberger (1976) where (1) is known as the asset’s gamma measure, $\gamma_i$, and defined (by analogy with the asset’s beta) as the ratio of the co-skewness of the asset’s return with the market to the market’s skewness $\gamma_i$. Just as $\beta_i$ represents the marginal contribution of the ith risky asset to the variance of the market portfolio $\sigma^2$, the asset’s marginal contribution to the portfolio skewness is captured by $\gamma_i$. Thus adding an asset with negative coskewness to a portfolio reduces the total skewness of the portfolio.
Summary statistics for the various risk factors considered are set out in Table 1, Panel A.

<table>
<thead>
<tr>
<th>Panel A: Summary statistics</th>
<th>Market</th>
<th>SMB</th>
<th>HML</th>
<th>Coskewness</th>
<th>Cokurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annualized mean return</td>
<td>0.0437</td>
<td>0.0286</td>
<td>0.0327</td>
<td>0.0049</td>
<td>0.0439</td>
</tr>
<tr>
<td>(0.11)</td>
<td>(0.17)</td>
<td>(0.23)</td>
<td>(0.87)</td>
<td>(0.18)</td>
<td></td>
</tr>
<tr>
<td>Annualized standard deviation</td>
<td>0.1539</td>
<td>0.1086</td>
<td>0.1212</td>
<td>0.1384</td>
<td>0.1636</td>
</tr>
<tr>
<td>Reward-to-variability</td>
<td>0.28</td>
<td>0.27</td>
<td>0.27</td>
<td>0.04</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Table 1
Summary statistics and correlations for the risk factors

Market is the return on the S&P500 Composite index in excess of the Treasury bill rate. SMB and HML are size and value risk premia as defined by Fama and French (1993). Coskewness is the return differential between two equally-weighted portfolios of stocks with low and high systematic skewness, respectively. Systematic skewness is defined as in Kraus and Litzenberger (1976). Cokurtosis is the return differential between two equally-weighted portfolios of stocks with high and low systematic kurtosis, respectively. The p-values in parentheses based on heteroskedasticity and autocorrelation robust (Newey-West) standard errors are for testing the significance of the mean. Reward-to-variability is the ratio of the annualized mean to the annualized standard deviation. For the pairwise correlations, the 95% confidence intervals reported in brackets are computed using Fisher's logarithmic transformation.

The prices of these systematic risks have signs that are consistent with risk aversion. Investors command higher returns on portfolios that have high market risk, small size, high book/market value, low skewness and high kurtosis. On average, the risk premium on the market is 4.37%, small capitalization stocks outperform large capitalization stocks by 2.86% while value stocks beat growth stocks by 3.27%. Portfolios with low systematic skewness outperform highly skewed portfolios by 0.49% on average, while investors require a premium of 4.39% for exposure to high (as opposed to low) levels of cokurtosis. The volatility of the risk factor-mimicking portfolios ranges from 10.66% for SMB to 16.36% for cokurtosis.

Table 1, Panel B, reports the pairwise correlations of the risk factors together with their 95% confidence intervals based on Fisher's transformation. The correlation of the cokurtosis factor with the SMB, market and coskewness factors ranges from 0.50 to 0.73 in absolute terms. The 95% confidence intervals suggest that these correlations are insignificantly different from 0.50, 0.60 and 0.75, respectively. Moreover, the $R^2$ of a regression of kurtosis risk on all other risk factors is 67.9%. Hence, the cokurtosis factor is excluded in Section 5 to address potential multicollinearity problems.

2.3. Asset-pricing models and hypotheses

Jensen’s alpha ($\alpha$) is the average return on a portfolio over and above that predicted by the CAPM. This measure can be obtained through the single index model (SIM hereafter)

$$R_t = \alpha + \beta_M R_{Mt} + \varepsilon_t,$$

where $R_t$ is either the momentum return or the excess return (over the risk-free rate proxied by Treasury bills) on the winner and loser portfolios and $R_{Mt}$ is the excess return on the S&P500 Composite index. The error term $\varepsilon_t$ represents idiosyncratic shocks with zero mean $E(\varepsilon_t) = 0$ and variance $E(\varepsilon_t^2) = \sigma^2 > 0$. The latter captures non-systematic risk.
Fama and French (1993) extend the SIM by including firm-based factors as candidates for systematic risk. They argue that small market-capitalization portfolios or high book/market value portfolios are relatively less desirable and hence, should command a risk premium. This motivates the three-factor Fama and French model (hereafter FFM)

\[ R_t = \alpha + \beta_M R_{M_t} + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \epsilon_t \]

where \( SMB_t \) (small minus big) is the return on a portfolio of small stocks in excess of that on a portfolio of large stocks and \( HML_t \) (high minus low) is the return on a portfolio of stocks with high book/market value in excess of that on a low book/market value portfolio.

As noted above, one of our goals is to assess whether departures from normality command a risk premium. For this purpose, we augment the SIM and FFM as follows

\[ R_t = \alpha + \beta_M R_{M_t} + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \beta_{SK} SK_t + \beta_{KU} KU_t + \epsilon_t \]

where \( SK_t \) and \( KU_t \) denote the returns of skewness- and kurtosis-mimicking portfolios.

This framework facilitates tests of two hypotheses (referred to as \( H_1 \) and \( H_2 \)) that relate to the idea that momentum profits are partly a compensation for non-normality in equity returns. First, the skewness and/or kurtosis risks of the winners should exceed those of the losers. Equivalently, the sensitivities of the momentum returns to the price of skewness and/or kurtosis should be positive (\( H_1 \)). Second, the alpha implied by the augmented SIM and FFM should be smaller than that implied by the baseline SIM and FFM, respectively (\( H_2 \)).

3. PRELIMINARY DATA ANALYSIS

Table 2 sets out summary statistics for the monthly returns of the winner, loser and momentum portfolios. Rows and columns signify ranking and holding periods, respectively.

<table>
<thead>
<tr>
<th></th>
<th>Holding period of 3 months</th>
<th></th>
<th>Holding period of 6 months</th>
<th></th>
<th>Holding period of 12 months</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Winner</td>
<td>Lower</td>
<td>Momentum</td>
<td>Winner</td>
<td>Lower</td>
<td>Momentum</td>
</tr>
<tr>
<td>Panel A: Ranking period of 3 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.1430</td>
<td>0.0996</td>
<td>0.0343</td>
<td>0.1390</td>
<td>0.0469</td>
<td>0.0967</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.1550</td>
<td>0.2251</td>
<td>0.1963</td>
<td>0.1905</td>
<td>0.2145</td>
<td>0.1227</td>
</tr>
<tr>
<td>Reward/variability</td>
<td>0.73</td>
<td>0.27</td>
<td>0.00</td>
<td>0.57</td>
<td>0.22</td>
<td>0.71</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.00</td>
<td>0.46</td>
<td>-0.29</td>
<td>-1.30</td>
<td>0.47</td>
<td>-0.46</td>
</tr>
<tr>
<td>Kurtosis-3</td>
<td>3.28</td>
<td>2.12</td>
<td>2.84</td>
<td>3.45</td>
<td>1.65</td>
<td>2.88</td>
</tr>
<tr>
<td>JB</td>
<td>0.39</td>
<td>1.10</td>
<td>1.51</td>
<td>0.19</td>
<td>0.90</td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Panel B: Ranking period of 6 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.1787</td>
<td>0.0941</td>
<td>0.1274</td>
<td>0.1916</td>
<td>0.0577</td>
<td>0.1413</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.2017</td>
<td>0.2215</td>
<td>0.1532</td>
<td>0.2121</td>
<td>0.2188</td>
<td>0.1455</td>
</tr>
<tr>
<td>Reward/variability</td>
<td>-0.89</td>
<td>0.16</td>
<td>0.91</td>
<td>-0.86</td>
<td>0.17</td>
<td>0.98</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.95</td>
<td>-0.30</td>
<td>-0.44</td>
<td>-0.86</td>
<td>-0.17</td>
<td>-0.36</td>
</tr>
<tr>
<td>Kurtosis-3</td>
<td>3.33</td>
<td>1.66</td>
<td>1.59</td>
<td>3.00</td>
<td>1.63</td>
<td>1.77</td>
</tr>
<tr>
<td>JB</td>
<td>0.09</td>
<td>0.28</td>
<td>0.93</td>
<td>0.15</td>
<td>0.27</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Panel C: Ranking period of 12 months</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.1679</td>
<td>0.0267</td>
<td>0.1612</td>
<td>0.1904</td>
<td>0.0406</td>
<td>0.1396</td>
</tr>
<tr>
<td>StdDev</td>
<td>0.30</td>
<td>0.58</td>
<td>0.60</td>
<td>0.03</td>
<td>0.43</td>
<td>0.03</td>
</tr>
<tr>
<td>Reward/variability</td>
<td>0.08</td>
<td>0.12</td>
<td>0.04</td>
<td>0.02</td>
<td>0.18</td>
<td>0.83</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.88</td>
<td>-0.16</td>
<td>-0.43</td>
<td>-0.74</td>
<td>-0.17</td>
<td>-0.24</td>
</tr>
<tr>
<td>Kurtosis-3</td>
<td>2.07</td>
<td>1.86</td>
<td>0.62</td>
<td>2.29</td>
<td>1.57</td>
<td>0.84</td>
</tr>
<tr>
<td>JB</td>
<td>0.54</td>
<td>4.63</td>
<td>13.35</td>
<td>0.00</td>
<td>5.41</td>
<td>12.50</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 2
Summary statistics of the returns on the winner, loser and momentum portfolios
The table reports summary statistics for the returns on the winner, loser and momentum portfolios. Winner (Loser) is an equally-weighted overlapping portfolio made of stocks that performed the best (worst) over a given ranking period. Momentum is a zero-cost portfolio that is long the winner portfolio and short the loser portfolio. Mean and standard deviation are annualized. Reward/variability is the ratio of the mean to the standard deviation. The \( p \)-values in brackets for the significance of the mean returns are based on heteroskedasticity and autocorrelation robust (Newey-West) standard errors. \( p \)-values are also reported for the significance of the skewness, excess kurtosis and the Jarque-Bera (JB) normality statistics.

The results confirm the presence of strong continuation in short term returns. The winner portfolios outperform the losers over horizons that range from 3 to 12 months. The momentum (winner minus loser) strategy produces positive returns which are strongly significant at better than the 1% level. The annualized mean momentum returns range from a low of 6.34% for the 12-12 momentum strategy to a high of 16.12% for the 12-3 strategy with an overall mean of 11.5%. Table 2 also reports the standard deviation of the portfolio returns. The winners exhibit smaller unconditional volatility than the losers, in line with the evidence in Harvey and Siddique (2000). This result is at odds with risk-aversion and portfolio theory and indicates that the returns to momentum trading are not a compensation for high volatility. A reward-to-variability ratio of 0.99, the 6-6 momentum strategy is the most profitable whereas the 12-12 strategy stands at the other end with a reward/variability ratio of 0.41. For comparison, the Sharpe ratio of the S&P500 index over the period is 0.29.

We test for the significance of the 3rd and 4th moments of the return distribution as given by the coefficients of skewness and (excess) kurtosis, respectively. The \( p \)-values reported build on the result that the skewness and kurtosis estimators follow \( N(0, 6/T) \) and \( N(3, 24/T) \) distributions, respectively, under the normality null for large samples (Stuart and Ord, 1987). There is strong evidence at the 1% level of negative skewness in all winner portfolios. In contrast, the loser portfolios on the 6-6, 12-3, 12-6, and 12-12 strategies have zero skewness. In the remaining cases, the return distribution of the winners is more negatively skewed than that of the losers in line with the evidence in Harvey and Siddique (2000). The upshot is that the distribution of the zero-cost (winner minus loser) momentum portfolios is significantly negatively skewed in 6 out of 9 instances: 3-3, 3-6, 6-3, 6-6, 12-3 and 12-6 strategies. The momentum returns are insignificantly skewed in two cases only (6-12 and 12-12) and significantly positively skewed in one instance (3-12).

The returns of both winners and losers have more mass in the tails than would be predicted by a normal distribution. Interestingly, in line with risk aversion and rational pricing in efficient markets, the winners stand out as relatively more leptokurtic than the losers (with only one exception, the 12-12 strategy). As a result, the excess kurtosis of the zero-cost momentum portfolios is significantly positive at the 1% level. The Jarque-Bera (JB) test statistic indicates that the distribution of winner, loser and momentum portfolios systematically departs from normality. In all but one strategy (12-12), the departure from normality is more substantial for the winners than for the losers. Figure 1 plots the distribution of the momentum returns for the 3-6 strategy.

![Figure 1](image)

*Zero-cost winner minus loser portfolio with a 3 month ranking period and a 6 month holding period*

The upper figure plots the
mean and actual returns on the momentum strategy over time. The lower figure plots the cumulative distribution function of momentum returns and that of normally distributed returns.

The top graph represents the monthly returns alongside the mean return. It provides prima facie evidence that the unconditional volatility of returns is time varying. The lower graph represents the cumulative distribution function (cdf) of the momentum returns alongside that of normally distributed returns. The null hypothesis that the empirical cdf of momentum returns is insignificantly different from a normal cdf is strongly rejected by the (finite-sample corrected) Kolmogorov-Smirnov test statistic at 1.89 with a \( p \)-value of 0.001\(^8\).

In the light of these findings, we conjecture that the profitability of momentum strategies may partly stem from the higher risks that winners bear relative to losers in terms of systematic skewness and kurtosis. Put differently, the abnormal returns of momentum strategies may reflect, to some extent, a market compensation for systematic risks not captured by conventional asset pricing models. This conjecture is investigated below.

4. MEASURING AND ANALYSING RISK-ADJUSTED PERFORMANCE

In this section, we first gauge the abnormal performance of momentum strategies by means of the baseline SIM and FFM frameworks outlined in Section 2. Next, we investigate whether the momentum effect is, at least partly, a compensation for the systematic risks associated with low skewness and high kurtosis. For this purpose, the SIM and FFM regressions are augmented with systematic skewness and kurtosis risk factors.

4.1. Market models: baseline and augmented with non-normality risks

The OLS estimates of the SIM and non-normality augmented SIM are set out in Table 3.

### Table 3 Baseline and augmented single index models

<table>
<thead>
<tr>
<th>Holding period of 3 months</th>
<th>Holding period of 6 months</th>
<th>Holding period of 12 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha )</td>
<td>( \beta_M )</td>
<td>( \beta_{SK} )</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>( \beta_M )</td>
<td>( \beta_{SK} )</td>
</tr>
<tr>
<td>( \beta_M )</td>
<td>( \beta_{SK} )</td>
<td>( \beta_{KU} )</td>
</tr>
</tbody>
</table>

\( \beta_{KU} \) is the abnormal performance measure. Michael is the portfolio returns to the market excess returns, the skewness- and kurtosis-mimicking portfolio returns, respectively. Winner (Loser) is an equally-weighted portfolio made of stocks that performed the best (worst) over the ranking period. Momentum is a portfolio that is long the winner portfolio and short the loser portfolio. \( R^2 \) is the adjusted coefficient of determination. White’s heteroscedasticity robust \( t \)-statistics are reported in parentheses.

\(^8\) Empirical distribution function tests are more robust than classical \( \chi^2 \) tests such as the JB test. The results for all other strategies (available upon request) are similar.
The residuals of each regression are subjected to: (a) White’s general test for heteroskedasticity and Engle’s ARCH test, (b) Breusch-Godfrey LM and Ljung-Box Q tests for autocorrelation. The former two tests provide strong evidence of time dependence in the conditional volatility of returns whereas the evidence of residual autocorrelation in the conditional mean is weak. Hence, the reported t-statistics are based on White’s heteroskedasticity-consistent variance estimator rather than on the OLS variance estimator.

The baseline SIM uniformly suggests that the market risk of the winners does not differ from that of the losers, the only exception being the 3-3 strategy for which the market beta of the momentum portfolio is negative at the 1% level. The alphas in the momentum return regressions are significantly positive for all strategies which, along with the insignificant market betas and low goodness-of-fit measures (adjusted $R^2$), suggests that momentum profitability is not a compensation for market risk. This finding is in line with the evidence in Jegadeesh and Titman (1993, 2001). The annualized alpha ranges from a low of 6.22% for the 12-12 strategy to a high of 16.43% for the 12-3 strategy. Buying recent winners and selling recent losers yields an alpha of 11.76% on average across the 9 strategies.

When we augment the SIM with the systematic skewness and kurtosis risk factors, as in equation (5), the results are rather encouraging. As conjectured ($H_3$), the skewness beta is larger for the winners than for the losers uniformly across all strategies. This indicates that investors holding winner portfolios expect a higher return to compensate them for exposure to a more negatively skewed distribution. Equivalently, investors are prepared to accept a lower return on loser portfolios ceteris paribus as these are less exposed to skewness risk. Most importantly, skewness significantly explains momentum returns generally at better than the 5% level in the 3-3, 6-3, 12-3, 12-6, and 12-12 strategies. In the remaining cases, the skewness beta of the momentum returns is also positive albeit insignificant.

The kurtosis beta is generally higher for the winner portfolios than for the loser portfolios but the difference is always insignificant as suggested by the t-ratio on $\beta_{KU}$ in the momentum regressions. Hence, skewness risk appears to play a more important role than kurtosis risk in explaining the zero-cost momentum returns. In other words, while momentum investors request compensation in the form of a higher return for exposure to negative skewness, departures from normality associated with kurtosis do not seem to be priced by the market.

The finding that systematic kurtosis plays an insignificant role in explaining momentum returns provides indirect support for the three-moment CAPM theory of Kraus and Litzenberger (1976). Preference for positive skewness is shown to aggregate (i.e., if all agents have a preference for positive skewness, then a representative investor will too), but this may not apply for kurtosis and other higher order moments. Kraus and Litzenberger show that when all investors exhibit non-increasing absolute risk aversion then, in a Pareto efficient allocation, the representative agent will also exhibit non-increasing absolute risk aversion. The latter implies that increases in the skewness of a portfolio are preferred. However, although individual investors will generally be averse to kurtosis, it does not necessarily follow that a representative investor will be averse to kurtosis.

According to the augmented SIM, the most successful strategy among the nine under study consists of selecting stocks based on their performance over the last 12 months and holding them for 3 months. This strategy yields an annualized alpha of 15.74%. As hypothesized ($H_3$), the average alpha of the augmented SIM at 11.32% is significantly smaller than that of the baseline SIM at 11.75% as suggested by a paired difference t-statistic of 6.13 with p-value of 0.00. Hence, skewness risk contributes to 42 basis points (bp) of the abnormal performance of relative-strength strategies. Moreover, the addition of skewness risk makes the SIM more competitive: the adjusted $R^2$ increases by 100bp on average and the difference is significant (t-statistic of 2.91 with a p-value of 0.01).

4.2. Fama-French models: baseline and augmented with non-normality risks
The estimation results for the baseline FFM, equation (4), and for the augmented FFM, equation (6), are set out in Table 4. The inference is based on White’s robust standard errors to accommodate heteroskedasticity given the evidence from both Engle’s ARCH and White’s heteroskedasticity tests on the residuals of (4) and (6).

---

11 - For instance, in the SIM of the 3-3 momentum returns, the White statistic for the ‘no heteroskedasticity’ null is 25.93 (p-value=0.00) whereas the Breusch-Godfrey statistic for ‘no autocorrelation up to 3 months’ is 6.134 (0.11). The results are qualitatively similar for all other strategies and are available upon request.
12 - The sample of differences $\{x_i, i=1, \ldots, n\}$ is used to make inferences about the mean of the population of differences. The test statistic $t=n(\bar{x}-\mu)/s$, where $\mu$ is the mean of $x_i$ and $s$ is the standard deviation, follows Student-t with $n-1$ degrees of freedom under $H_0:x=\mu$. Independence between the two samples is not required nor equality of variances.
According to the FFM, the market beta ($\beta_M$) of the winners is generally smaller than that of the losers. The difference is significant in 4 strategies (3-3, 3-6, 6-3 and 12-3). The stocks of both winner and loser portfolios are relatively small in size as evidenced by their significantly positive size beta ($\beta_{SMB}$) which is consistent with the findings in Lesmond et al. (2004). The momentum regressions indicate that the market capitalization of winners and losers does not differ significantly. Hence, size risk cannot explain the momentum effect. Winner portfolios appear to pick growth stocks (as suggested by their negative $\beta_{HML}$) whereas loser portfolios tend to hold value stocks (positive $\beta_{HML}$). Relatedly, the book/market beta of momentum returns is always negative and significant at the 5% level in virtually all cases which suggests that book/market risk is a determinant of momentum returns.

The explanatory power (adjusted $R^2$) of the FFM at 9.18% on average rises substantially relative to that of the SIM at 0.68% (cf. Table 3). However, the risk-adjusted performance measure of the momentum portfolios in the FFM remains significantly positive throughout the strategies which suggests, in line with the evidence in Fama and French (1996), Jegadeesh and Titman (2001) and Karolyi and Kho (2004), that the three-factor FFM is not successful in explaining the momentum effect. As the momentum portfolios have negative betas on $HML$, the mean abnormal performance of the FFM (at 13.6%) is higher than that of the SIM (at 11.75%). As in Table 3, the 12-3 strategy stands out as the most successful. Note also that, in line with Moskowitz and Grinblatt (1999) and Hong et al. (2000), the momentum effect appears to be driven mainly by the losers: the annualized alpha of loser portfolios ranges from -12.14% to -7.78% and is clearly significant at the 1% level in all cases whereas that of the winners, in the range [1.15%, 7.13%], is significant in 4 out of 9 cases only.

The explanatory power (adjusted $R^2$) of the FFM at 9.18% on average rises substantially relative to that of the SIM at 0.68% (cf. Table 3). However, the risk-adjusted performance measure of the momentum portfolios in the FFM remains significantly positive throughout the strategies which suggests, in line with the evidence in Fama and French (1996), Jegadeesh and Titman (2001) and Karolyi and Kho (2004), that the three-factor FFM is not successful in explaining the momentum effect. As the momentum portfolios have negative betas on $HML$, the mean abnormal performance of the FFM (at 13.6%) is higher than that of the SIM (at 11.75%). As in Table 3, the 12-3 strategy stands out as the most successful. Note also that, in line with Moskowitz and Grinblatt (1999) and Hong et al. (2000), the momentum effect appears to be driven mainly by the losers: the annualized alpha of loser portfolios ranges from -12.14% to -7.78% and is clearly significant at the 1% level in all cases whereas that of the winners, in the range [1.15%, 7.13%], is significant in 4 out of 9 cases only.

The augmented-FFM estimates suggest, in line with hypothesis $H_2$, that the skewness beta of the winners uniformly exceeds that of the losers and the difference is significant in 5 strategies (3-3, 6-3, 12-3, 12-6, 12-12) generally at better than the 10% level. In the remaining cases, the skewness risk of the momentum portfolios is also positive but insignificant. By contrast, the kurtosis beta is generally insignificant as in the SIM framework.
Interestingly, once non-normality risk is taken into account, the momentum profitability measure decreases uniformly for all strategies. In the five momentum regressions where the skewness beta is significant, the decrease in \( \alpha \) ranges from 61bp for the 6-3 strategy to 94bp for the 12-3 strategy. In line with hypothesis \( H_2 \), the average \( \alpha \) in the augmented FFM at 13% is significantly smaller than that in the baseline FFM at 13.6% as suggested by a paired difference t-statistic of 6.09 with p-value of 0.00. This suggests that skewness risk contributes to about 60bp of the annual abnormal performance of momentum strategies. Moreover, by including non-normality risks, the explanatory power of the asset-pricing model improves as borne out by a significant increase in the adjusted \( R^2 \) from 11.6% (FFM) to 14.1% (augmented FFM) for the five cases where \( \beta_{SK} \) is significant or from 9.2% to 10.8% overall (for the latter, the paired difference t-statistic is 4.17 with p-value 0.002). The upshot is that systematic skewness risk can explain a significant part of the momentum effect over and above the market, size and book/market risks. However, the still relatively high alpha (at 13% on average) in the augmented FFM indicates that liquidity risk, trading costs and behavioural biases may play also crucial roles in explaining the momentum puzzle.

5. CONDITIONAL RISKS AND ABNORMAL PERFORMANCE

The above analysis is now re-conducted in a more general framework to deal with three specification problems. First, in Section 5.1., we explicitly tackle the time variation in the idiosyncratic risk of the momentum strategies. Second, in Section 5.2., we deal jointly with the latter and with time variation in the systematic risks and in the risk-adjusted performance. Third, in the light of the previous findings and in order to mitigate potential multicollinearity problems, the systematic kurtosis risk factor is excluded hereafter\(^{13}\).

5.1. Time dependence in the idiosyncratic risk of momentum portfolios

In a seminal paper, Pagan and Schwert (1990) compute recursive estimates of the variance of monthly stock returns and demonstrate how important it is to consider time variation in stock variance. In the present context, Engle’s LM test on the OLS residuals of the asset-pricing models (3) to (6) produces strong evidence of ARCH or volatility clustering effects\(^{14}\). Since the residual variance reflects non-systematic risk, the ARCH effects may stem from persistence in firm-specific uncertainty. In this context, the usual OLS estimates are still unbiased and consistent (albeit inefficient) and the OLS standard errors are incorrect. We seek to obtain efficiency gains in the risk-adjusted performance estimates by fitting two simultaneous equations, one to explain the momentum returns and another to explain the time-varying idiosyncratic risk. Formally, we estimate the following SIM specification

\[
R_t = \alpha + \beta_M R_{Mt} + \varepsilon_t, \quad \varepsilon_t = \epsilon_t \sigma_t, \quad \epsilon_t \sim \text{NID}(0, 1)
\]

and likewise for the FFM model and for the skewness-augmented SIM and FFM. The above Generalized ARCH (GARCH) variance equation states that the volatility of portfolio returns conditional on the market is driven by the history of the shocks to returns and by the history of the volatility itself. We adopt \( p=q=1 \) and use either maximum likelihood (ML) estimation if the model’s standardized errors, \( e_t \) are normal or the quasi-ML (QML) approach otherwise. Two model-adequacy measures based on the standardized residuals are typically used for these dynamic models, the Jarque-Bera (JB) normality test and the Akaike Information Criterion (AIC). The AIC trades off model-fit against parsimony by adjusting the average log-likelihood function with a penalty for degrees of freedom.

Table 5 summarizes the results. In all models, Engle’s LM statistic suggests that the GARCH(1,1) specification succeeds in capturing the time variation in the conditional volatility of returns; namely, there is no neglected conditional heteroskedasticity in \( e_t \). There is no evidence of residual autocorrelation either\(^{15}\). Hence, the standardized errors are white noise.

---

13 - The kurtosis beta appears also insignificant in the robust asset-pricing models of Section 5 in line with our earlier findings. The significance of the systematic skewness beta is invariant to the inclusion/exclusion of the systematic kurtosis factor.

14 - For instance, in model (6) for the momentum returns of the 3-6, 12-3 and 6-6 strategies, the LM test statistic for ‘no ARCH up to order 3’ is 96.07 (p-value=0.00), 44.49 (p=0.00) and 87.20 (p=0.00), respectively. A complete set of results is available upon request.

15 - For instance, in model (6) with a GARCH term for the 3-6, 12-3 and 6-6 strategies the LM test statistic for ‘no ARCH up to order 3’ is 1.36 (p-value=0.71), 0.75 (p=0.86) and 1.44 (p=0.70), respectively; the Ljung-Box statistic for ‘no autocorrelation up to order 3’ in the 3-6, 12-3 and 6-6 strategies is 1.33 (p=0.72), 0.56 (p=0.91) and 0.89 (p=0.38), respectively.
In the augmented SIM, the skewness beta is positively signed and significant at better than the 5% level throughout which provides support for hypothesis $H_1$. On average, a 1% increase in the price of skewness risk leads to a 0.16% increase in momentum returns. Most importantly, the addition of systematic skewness risk to the SIM leads to a fall in the annualized $\alpha$ that ranges from 60bp (12-12 strategy) to 120bp (3-12 strategy). On average, skewness risk explains 90bp of the observed annual abnormal performance. The average $\alpha$ from the SIM-SK at 9.6% is significantly smaller than that from the SIM at 10.5% (paired difference $t$-statistic = 14.05; $p$-value = 0.00) in line with hypothesis $H_2$. The SIM-SK is also more competitive than the SIM as suggested by the JB and AIC statistics uniformly falling. Moreover, skewness risk is able to explain some of the momentum returns over and above the market, size and book/market risks as borne out by the significantly positive skewness beta in the FFM-SK at the 1% level in virtually all strategies. The latter provides further support for hypothesis $H_1$. Other things being equal, the estimates in Table 5 suggest that investors require an average increase in returns of 0.18% per 1% increase in the price of skewness risk. Interestingly, a comparison of the FFM and FFM-SK estimates reveals that, once skewness risk is accounted for, market risk ($\beta_M$) and, more particularly, size risk ($\beta_{SMB}$) are relatively less important. For the latter, clear instances are the 3-6, 6-3, 12-6 and 12-12 strategies. Hence, skewness risk may capture some of the economic risks associated with firm size. Multicollinearity would increase the uncertainty of the parameter estimates; however, the estimates of the FFM-SK could go in either direction relative to the FFM estimates. However, as Table 5 shows, the inclusion of skewness risk in the momentum regressions shifts the size beta in the same direction uniformly across all the strategies. On the other hand, the pervasive significance of the book/market beta ($\beta_{HML}$) corroborates our earlier finding that momentum strategies tend to pick up growth stocks.

Moreover, skewness risk is able to explain some of the momentum returns over and above the market, size and book/market risks as borne out by the significantly positive skewness beta in the FFM-SK at the 1% level in virtually all strategies. The latter provides further support for hypothesis $H_1$. Other things being equal, the estimates in Table 5 suggest that investors require an average increase in returns of 0.18% per 1% increase in the price of skewness risk. Interestingly, a comparison of the FFM and FFM-SK estimates reveals that, once skewness risk is accounted for, market risk ($\beta_M$) and, more particularly, size risk ($\beta_{SMB}$) are relatively less important. For the latter, clear instances are the 3-6, 6-3, 12-6 and 12-12 strategies. Hence, skewness risk may capture some of the economic risks associated with firm size. Multicollinearity would increase the uncertainty of the parameter estimates; however, the estimates of the FFM-SK could go in either direction relative to the FFM estimates. However, as Table 5 shows, the inclusion of skewness risk in the momentum regressions shifts the size beta in the same direction uniformly across all the strategies. On the other hand, the pervasive significance of the book/market beta ($\beta_{HML}$) corroborates our earlier finding that momentum strategies tend to pick up growth stocks.

Moreover, skewness risk is able to explain some of the momentum returns over and above the market, size and book/market risks as borne out by the significantly positive skewness beta in the FFM-SK at the 1% level in virtually all strategies. The latter provides further support for hypothesis $H_1$. Other things being equal, the estimates in Table 5 suggest that investors require an average increase in returns of 0.18% per 1% increase in the price of skewness risk. Interestingly, a comparison of the FFM and FFM-SK estimates reveals that, once skewness risk is accounted for, market risk ($\beta_M$) and, more particularly, size risk ($\beta_{SMB}$) are relatively less important. For the latter, clear instances are the 3-6, 6-3, 12-6 and 12-12 strategies. Hence, skewness risk may capture some of the economic risks associated with firm size. Multicollinearity would increase the uncertainty of the parameter estimates; however, the estimates of the FFM-SK could go in either direction relative to the FFM estimates. However, as Table 5 shows, the inclusion of skewness risk in the momentum regressions shifts the size beta in the same direction uniformly across all the strategies. On the other hand, the pervasive significance of the book/market beta ($\beta_{HML}$) corroborates our earlier finding that momentum strategies tend to pick up growth stocks.
The FFM-SK estimates indicate that the risk-adjusted profitability of all momentum strategies falls when skewness risk is accounted for. The decrease in \( \alpha \) ranges from 36bp (3-3 strategy) to 108bp (12-3 strategy) with a mean of 72bp. Using the most profitable strategy (12-3) as an example, the annualized alpha of FFM-SK is 16.3% whereas that of FFM is 17.4%. As conjectured in \( H_2 \), the mean alpha from the FFM-SK at 11.4% is significantly smaller than that of FFM at 12.2% (paired difference t-statistic = 9.3). Moreover, the fall in alpha comes accompanied by more favourable JB and AIC model diagnostics.

However, the alpha measures from the present GARCH framework imply that, while the prices of market, book/market value and skewness risks have a role to play, they still fail to explain a large proportion (above 11%) of the momentum profitability.

### 5.2. Time dependence in abnormal performance and systematic risks

The asset-pricing models used thus far assume constant abnormal performance and systematic risks. However, the momentum asset allocation is dynamic by definition and therefore the systematic risks are bound to change over time. Moreover, the inexorable succession of economic expansions and recessions is also likely to change the systematic risks of stocks. In turn, non-constant risks lead to rational variation in required returns.

We adopt Harvey's (1989) conditioning framework to introduce time variation in the risk exposures and in the risk-adjusted performance of the momentum strategies. Accordingly, let \( Z_t \) denote a variable that captures changing economic conditions and dynamic asset allocations. The time variation in alpha is modeled linearly as \( \alpha_t = \alpha_0 + \alpha_1 Z_{t-1} \). The time variation in the market risk is modeled as \( \beta_{M,t} = \beta_{0,M} + \beta_{1,M} Z_{t-1} \) and so forth. Following Christopherson et al. (1998), we estimate the conditional-FFM equation

\[
R_t = \alpha_0 + \alpha_1 Z_{t-1} + \beta_{0,M} R_{M,t} + \beta_{1,M} Z_{t-1} R_{M,t} + \beta_{0,SMB} SMB_t + \beta_{1,SMB} Z_{t-1} SMB_t + \beta_{0,HML} HML_t + \beta_{1,HML} Z_{t-1} HML_t + \varepsilon_t
\]

(7)

where \( Z_{t-1} \) is demeaned so that \( \alpha_0 \) gives the expected abnormal performance and \( \beta_{0,M} \beta_{0,SMB} \) and \( \beta_{0,HML} \) give the expected systematic risks. It follows that \( \beta_{1,M} Z_{t-1} \) captures the time variation in market risk that is related to changing economic conditions and dynamic asset allocations. Likewise, \( \alpha_1 Z_{t-1} \) captures the variation through time in abnormal performance. Since the foregoing analysis has revealed dynamics in idiosyncratic risk, a GARCH equation is added to the above conditional FFM and these two simultaneous equations are estimated by (Q)ML.

We focus the discussion on (7) and a skewness-augmented version.

Following the literature, the candidates for \( Z \) initially considered are the monthly default spread, dividend yield, Treasury bill and term structure (Chordia and Shivakumar, 2002; Wu, 2002). However, on the basis of a multicollinearity analysis, only default spread and term structure are included in the final conditioning vector, \( Z_t = (DS_t, TS_t)^T \). Thus we have \( \alpha_t = \alpha_0 + \alpha_1 DS_{t-1} + \alpha_2 TS_{t-1} \) for alpha, \( \beta_{SK,t} = \beta_{0,SK} + \beta_{1,SK} DS_{t-1} + \beta_{2,SK} TS_{t-1} \) for the skewness beta and likewise for the remaining betas. Table 6 reports the estimates of the conditional-FFM without and with skewness risk. The test statistics reported are: (i) t-ratios for the individual significance of \( \alpha_0 \) and \( \beta_{0,SK} \) and (ii) Wald statistics for the joint hypothesis of no time variation in the parameters of the asset-pricing model. For instance, the null joint hypothesis for alpha is \( H_0: \alpha = 0 \) with \( \alpha = (\alpha_1, \alpha_2)^T \) and a rejection suggests that \( Z_{t-1} \) has predictive power over abnormal performance. We proceed likewise with the systematic risks.

---

18 - The mean adjusted R² of FFM-SK exceeds that of FFM by a significant 1.4% (t-statistic = 2.96).
19 - Engle's LM test yields clear evidence of ARCH effects in the residuals of equation (7).
20 - DS is the difference in yields between Moody’s Baa and Aaa rated corporate bonds and TS is the difference in the yield on US T-bonds with at least 10-year maturity and the 3-month T-bill rate. The data are obtained from Datastream.
There is evidence of instability in the market, book/market value and skewness risks. The FFM-SK model yields significant Wald statistics for the market beta in 7 strategies, for the HML beta in 7 strategies and for the skewness beta in 5 instances. Evidence of time variation in alpha is found in 4 out of the 9 strategies. In sharp contrast, the systematic risk associated with size (SMB beta) appears not only constant but insignificantly different from zero. Overall, our findings suggest that the systematic risks and abnormal performance of momentum portfolios are not constant over time. This strongly motivates the use of a conditional-FFM model to re-assess the evidence on the profitability of relative-strength strategies and the importance of systematic skewness risk.

The skewness beta ($\beta_{0,SK}$) is significantly positive throughout the strategies in line with hypothesis $H_1$. A comparison of the pricing models with and without skewness risk is quite revealing. First, the model-diagnostics JB and AIC improve (fall) in virtually all cases when systematic skewness risk is controlled for. Second, the expected abnormal performance ($\alpha_0$) falls when skewness is accounted for. For instance, in annualized terms the decrease amounts to more than 100bp for the 6-12 and 12-3 strategies. As conjectured ($H_2$), the addition of skewness risk reduces the average alpha of the conditional FFM from 11.9% to 11.2% ($t$-statistic $= 4.71$) which suggests that skewness risk contributes significantly to 70bp of the annual abnormal returns. Compounding $1 over the sample period at the annualized alpha of the baseline and augmented conditional-FFM yields $23.59$ and $20.20$, respectively. This indicates that an asset manager who misleadingly took credit for negative skewness over the sample period is attributing $3.39 per dollar invested to his stock picking ability, while this $3.39 actually was a reward for systematic skewness. This highlights the harmful impact that negative coskewness has on portfolio performance.

---

**Table 6**

**Momentum return models with conditional performance and conditional risks**
The table reports results from conditional momentum regressions with time-varying idiosyncratic risk of GARCH(1,1) type. $\alpha_0 = E(\alpha_t)$ is the expected abnormal performance. $\beta_{0,SK} = E(\beta_{SK,t})$ is the expected skewness risk. $t$-statistics are reported in parenthesis. $p(\alpha_0)$, $p(\beta_{M}=0)$, $p(\beta_{SMB}=0)$, $p(\beta_{HML}=0)$, and $p(\beta_{SK}=0)$ are the p-values of Wald statistics to test the null hypothesis of no time variation in $\alpha_0$, $\beta_{M}$, $\beta_{SMB}$, $\beta_{HML}$ and $\beta_{SK}$, respectively, from a $x^2$ distribution with 2 degrees of freedom. JB is the Jarque-Bera statistic with p-values in brackets. The model with the lowest Akaike Information Criterion (AIC) is superior.

<table>
<thead>
<tr>
<th>Holding = 3 months</th>
<th>FFM</th>
<th>FFM-SK</th>
<th>Holding = 6 months</th>
<th>FFM</th>
<th>FFM-SK</th>
<th>Holding = 12 months</th>
<th>FFM</th>
<th>FFM-SK</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_0$</td>
<td>0.0075</td>
<td>0.0074</td>
<td>0.0073</td>
<td>0.0070</td>
<td>0.0067</td>
<td>0.0062</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3.9765)</td>
<td>(3.8857)</td>
<td>(4.7724)</td>
<td>(4.6612)</td>
<td>(5.5278)</td>
<td>(5.3894)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{SK}$</td>
<td>—</td>
<td>0.1525</td>
<td>—</td>
<td>0.1564</td>
<td>—</td>
<td>0.1221</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2.0607)</td>
<td>(2.0603)</td>
<td>(2.0607)</td>
<td>(2.0603)</td>
<td>(2.0603)</td>
<td>(2.0603)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p(\alpha_0)$</td>
<td>0.5778</td>
<td>0.4516</td>
<td>0.3119</td>
<td>0.1614</td>
<td>0.0820</td>
<td>0.0719</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p(\beta_{M}=0)$</td>
<td>0.8050</td>
<td>0.5595</td>
<td>0.5843</td>
<td>0.0454</td>
<td>0.3465</td>
<td>0.0567</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p(\beta_{SMB}=0)$</td>
<td>0.8050</td>
<td>0.6925</td>
<td>0.0458</td>
<td>0.1327</td>
<td>0.1471</td>
<td>0.2516</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p(\beta_{HML}=0)$</td>
<td>0.3560</td>
<td>0.2147</td>
<td>0.0766</td>
<td>0.0104</td>
<td>0.0023</td>
<td>0.0002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$p(\beta_{SK}=0)$</td>
<td>—</td>
<td>0.0321</td>
<td>—</td>
<td>0.0161</td>
<td>—</td>
<td>0.2675</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[0.0000]</td>
<td>[0.0000]</td>
<td>[0.0016]</td>
<td>[0.0064]</td>
<td>[0.0000]</td>
<td>[0.0000]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

**Panel B: Ranking period of 6 months**

| $\alpha_0$         | 0.0115 | 0.0117 | 0.0118 | 0.0114 | 0.0056 | 0.0068 |
| $\beta_{SK}$       | —     | 0.1914 | —     | 0.2125 | —     | 0.2064 |
| (2.7309)           | (3.8567) | (2.7309) | (3.8567) | (2.7309) | (3.8567) |
| $p(\alpha_0)$      | 0.5543 | 0.2716 | 0.0591 | 0.0836 | 0.3191 | 0.1334 |
| $p(\beta_{M}=0)$   | 0.5501 | 0.0675 | 0.2214 | 0.0339 | 0.2522 | 0.1334 |
| $p(\beta_{SMB}=0)$ | 0.0932 | 0.2653 | 0.0572 | 0.1630 | 0.6431 | 0.7507 |
| $p(\beta_{HML}=0)$ | 0.2244 | 0.0135 | 0.0630 | 0.0005 | 0.6783 | 0.7111 |
| $p(\beta_{SK}=0)$  | —     | 0.0835 | —     | 0.0018 | —     | 0.1060 |
| JB                 | 6.3476 | 5.1902 | 6.3757 | 5.3018 | 36.6484 | 68.4273 |
| [0.0418]           | [0.0746] | [0.0413] | [0.1796] | [0.0000] | [0.0000] |

**Panel C: Ranking period of 12 months**

| $\alpha_0$         | 0.0149 | 0.0139 | 0.0129 | 0.0125 | 0.0062 | 0.0057 |
| (6.9406)           | (6.5221) | (5.3063) | (5.2137) | (3.2097) | (2.6619) |
| $\beta_{SK}$       | —     | 0.3191 | —     | 0.2645 | —     | 0.2245 |
| $p(\alpha_0)$      | 0.4088 | 0.0617 | 0.1862 | 0.0474 | 0.0607 | 0.1697 |
| $p(\beta_{M}=0)$   | 0.0106 | 0.0009 | 0.0534 | 0.0405 | 0.0339 | 0.0005 |
| $p(\beta_{SMB}=0)$ | 0.3090 | 0.6157 | 0.5542 | 0.4974 | 0.0767 | 0.3275 |
| $p(\beta_{HML}=0)$ | 0.2733 | 0.1949 | 0.0210 | 0.0696 | 0.0028 | 0.0000 |
| $p(\beta_{SK}=0)$  | —     | 0.9272 | —     | 0.3698 | —     | 0.0000 |
| JB                 | 0.7827 | 0.594 | 4.8288 | 9.8075 | 7.1914 | 5.6739 |
| [0.6762]           | [0.7430] | [0.0884] | [0.0074] | [0.0236] | [0.0586] |

---

21 - Wu (2002) finds little evidence of time variation in market risk but strong evidence of time variation in the SMB and HML betas. Our results are invariant to assuming constant idiosyncratic risk, $E(\varepsilon_t^2) = \sigma^2$, so the contrast may stem from the different conditioning sets.
One expects rational risk-averse investors and market timers to tilt their asset allocation towards the best performing asset class; namely, towards relatively safer assets in down markets and relatively riskier assets in up markets. In this light, the winner (loser) portfolios ought to be more (less) risky in up markets than in down markets. Put differently, during economic expansions the winner (loser) portfolios should comprise stocks with relatively higher (lower) market risk, smaller (larger) size, higher (lower) book/market value and lower (higher) skewness than during recessions.

The hypothesis that the asset allocation of momentum strategies changes over the business cycle can be tested by regressing the time-varying risk measures obtained from (7) on a constant and the year-on-year change in industrial production ($IP_t$) as a proxy for the business cycle. For instance, we estimate the regression model $\beta_{M,t} = \gamma + \delta_{M} IP_t + \epsilon_t$ where the variable to be explained is the time-varying market beta defined as $\beta_{M,t} = \beta_{M,0} + \beta_{M,1} Z_{t-1}$. The sign of the slope coefficient ($\delta_{M}$) indicates the direction in which market risk changes over the business cycle. We proceed likewise with the other systematic risks. Although the Wald tests in Table 6 suggest insignificant time variation in the betas for some of the strategies, all cases are considered in the ensuing analysis for completeness. Table 7 reports the results.

The slope coefficient of the market risk ($\beta_{M,t}$) and skewness risk ($\beta_{SK,t}$) regressions is positive for the winners and negative for the losers in virtually all the strategies. This suggests that during up markets, as industrial production rises, the momentum strategy tends to pick winner (loser) stocks that have higher (lower) market and skewness risks than during down markets. Relatedly, the momentum return regressions suggest that in expansionary episodes the momentum portfolios consist of stocks which have higher systematic market risk and skewness risk than during recessionary periods. These findings are in line with the notion that risk averse investors and market timers tactically tilt their asset allocation towards relatively riskier (safer) assets in up (down) markets. Similarly, the momentum strategy recommends investing in the equities that consistently offer the best returns.

The evidence regarding the systematic risks associated with size ($\beta_{SMB,t}$) and book/market value ($\beta_{HML,t}$) is mixed and no clear pattern emerges for winners and losers. If anything, the results uncover a tendency for winner, loser and momentum portfolios to pick companies with relatively large size and higher book/market values during up markets. Moreover, the regression fit is notably better for the market and skewness risks than for the SMB.
and HML risks. For instance, the mean $R^2$ over strategies for the winner regressions is 14% for both the market and skewness betas whereas it falls to 4% and 3% for the SMB and HML betas, respectively. The mean $R^2$ for the momentum regressions is 23% and 20% for the market and skewness betas, respectively, whereas it roughly halves to 10% and 8% for the SMB and HML betas, respectively. In order to mitigate potential endogeneity biases we re-estimated the regressions using 1-month lagged industrial production changes instead. Although the $R^2$ somewhat improves, the results are qualitatively similar.

6. CONCLUSIONS
While the profitability of momentum strategies is rarely disputed, there is still controversy over the reasons for such abnormal returns. This paper investigates the conjecture that the latter are partly a compensation for non-normality in the distribution of momentum returns. We corroborate the findings in Harvey and Siddique (2000) that the return distribution of winner portfolios is more negatively skewed than that of the losers. The former also has higher excess kurtosis than the latter. Relatedly, the distribution of momentum returns shows significant negative skewness and leptokurtosis. Against this background, it seems natural to investigate the extent to which systematic non-normality risks explain the momentum effect. For this purpose, we adopt a systematic skewness (or coskewness) definition that builds on the three-moment CAPM theory of Kraus and Litzenberger (1976) and we extend it also to the fourth moment. Building on the methodology in Fama and French (1993), we form skewness- and kurtosis-mimicking portfolio returns which are subsequently included as additional risk factors in the single index model and the Fama-French model.

The results support the notion that momentum profits are partly a compensation for exposure to undesired systematic skewness. This finding is pervasive across the nine momentum strategies that arise from combining 3-, 6-, and 12-month ranking and holding periods. Overall, skewness risk contributes significantly to about 70bp of the annual abnormal returns of momentum strategies. The main conclusion is that momentum strategies involve more risks than hitherto implied by conventional asset-pricing models.

The paper documents conditional heteroskedasticity in the residuals of the single index model and the Fama-French model which indicates that there is time dependence in the idiosyncratic risk of momentum portfolios. Interestingly also, the risk exposures of momentum portfolios appear to change over time as market-wide economic conditions vary. More specifically, the momentum portfolios tend to be riskier during economic expansions as they are long (short) stocks with relatively higher (lower) market beta and negative (positive) skewness than during recessions. These findings are consistent with the tactical asset allocation that risk averse investors and market timers would naturally adopt.

What do we learn from these findings? Momentum profits are partly a compensation for exposure to higher-order moments such as skewness in asset returns. But altogether the systematic market, size, book/market and skewness risks are unable to fully rationalize the momentum effect. Hence, lack of liquidity and/or transaction costs cannot be ruled out as additional explanations. Researchers of a behavioural persuasion can take comfort in the fact that a large proportion of the momentum returns remains unexplained.

Interesting avenues remain open for further research. It is possible that negative skewness in the return distribution of winner portfolios arises from the fact that the underlying companies disclose bad news within a 12-month window after portfolio formation. An event study could shed light on this hypothesis. A generalization of the asset pricing models here considered that incorporates transaction costs could also prove fruitful.
REFERENCES


• Brennan M., 1993, "Agency and asset pricing". Finance; 1; 6-93

• Christ-David R, Chaudhry M., 2001, "Coskewness and cokurtosis in futures markets". Journal of Empirical Finance; 8; 55-81

• Chordia T, Shivakumar L., 2002, "Momentum, business cycle, and time-varying expected returns". Journal of Finance; 57; 985-1018


• Chung P Y, Johnson H, Schill M J., 2005, "Asset pricing when returns are nonnormal: Fama-French factors vs. higher-order systematic co-moments". Journal of Business


• Fama E F, French K R., 1996, "Multifactor explanations of asset pricing anomalies". Journal of Finance; 51; 55-84

• Golec J, Tamarkin M., 1998, "Bettors love skewness, not risk, at the horse track". Journal of Political Economy; 106; 205-225


• Harvey C. 1995 "Predictable risk and returns in emerging market". Review of Financial Studies; 8; 773-816

• Harvey C R, Siddique A., 2000, "Conditional skewness in asset pricing tests". Journal of Finance; 55; 1263-1295

• Hong H, Lim T, Stein J., 2000, "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies". Journal of Finance; 55; 265-295

• Hong H, Stein J., 1999, "A unified theory of underreaction, momentum trading and overreaction in asset markets". Journal of Finance; 54; 2143-2184

• Jegadeesh N, Titman S., 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency". Journal of Finance; 48; 65-91
  *Journal of Finance*; 56; 699-720


• Kraus A, Litzenberger R., 1976, "Skewness preference and the valuation of risk asset". *Journal of Finance*; 31; 1085-1100


• Moskowitz T J, Grinblatt M., 1999, "Do industries explain momentum?" *Journal of Finance*; 54; 1249-1290


• Sadka R., 2003, "Momentum, liquidity risk, and limits to arbitrage", *Working paper Northwestern University*


• Wu X., 2002, "A conditional multifactor model of return momentum". *Journal of Banking and Finance*; 26; 1675-1696