September 5, 2012 Comments Welcome

# Tail Risk in Momentum Strategy Returns

Kent Daniel, Ravi Jagannathan and Soohun Kim*

**Abstract**

Price momentum strategies have historically generated high positive returns with little systematic risk. However, these strategies also experience infrequent but severe losses. During 13 of the 978 months in our 1929-2010 sample, losses to a US-equity momentum strategy exceed 20 percent per month. We demonstrate that a hidden Markov model in which the market moves between latent "*turbulent*" and "*calm*" states in a systematic stochastic manner captures these high-loss episodes. The turbulent state is infrequent in our sample: the probability that the hidden state is turbulent is greater than one-half in only 20% of the months. Yet in each of the 13 severe loss months, the *ex-ante* probability that the hidden state is turbulent exceeds 70 percent. This strong forecastability accentuates the price momentum puzzle; a conditional momentum strategy that moves to the risk-free asset when the *ex-ante* probability of the turbulent state is high exhibits dramatically better performance than the unconditional momentum strategy.

*Columbia University, Graduate School of Business (Daniel) and Northwestern University, Kellogg School (Jagannathan and Kim). This paper originally appeared under the title "Risky Cycles in Momentum Returns". We thank Raul Chhabbra, Randolph B. Cohen, Zhi Da, Gangadhar Darbha, Francis Diebold, Robert Engel, Bryan T. Kelly, Robert Korajczyk, Jonathan Parker, Prasanna Tantri, the participants of seminars at Northwestern University, the Indian School of Business, the Securities and Exchange Board of India, Shanghai Advanced Institute for Finance, Shanghai Jiao Tong University, Fordham University, and Nomura Securities, and the participants of the Fourth Annual Conference of the Society for Financial Econometrics, the Fifth Annual Triple Crown Conference in Finance, and the 2012 WFA meeting for many helpful comments and suggestions. We alone are responsible for any errors and omissions.

# 1 Introduction

Relative strength strategies, also known as price momentum strategies, have been and continue to be popular among traders. Quantitative strategies used by active money managers often rely on some form of momentum.[1] Even those who use fundamental analysis appear to incorporate momentum into their trading decisions.[2]

Price momentum in stocks can be described as the tendency of those stocks with the highest (lowest) past return to subsequently outperform (underperform) the broader market. Price momentum strategies exploit this pattern by taking a long position in past winners and an equal short position in past losers. These strategies produce high abnormal returns on average, but also generate infrequent and large losses.

Over our sample period of July 1929 through December 2010, our baseline momentum strategy produces monthly excess returns with a mean of 1.12%, a standard deviation of 8%, and with little systematic risk as measured by the standard CAPM or the Fama and French (1993) three factor model: the estimated CAPM $\alpha$ of the strategy is 1.44%/month, and the Fama and French (1993) three factor $\alpha$ is 1.70%/month.[3] The maximum attainable sample monthly Sharpe Ratio increases from 0.15 for a portfolio of the three Fama and French (1993) factors to 0.28 when the three factors are augmented with the momentum strategy.

However, the momentum strategy returns have an empirical distribution that is both highly left skewed and significantly leptokurtic. There are thirteen months with losses exceeding 20%. In its worst month the strategy experiences a loss of 79%. Were momentum returns generated from an *i.i.d.* normal distribution with mean and variance equal to their sample counterparts, the probability of realizing a loss of more than 20% in thirteen or more months would be 0.04%, and the probability of incurring a loss of 79% or worse would be $9.95\times10^{-24}$.

[1] See Swaminathan (2010), who further estimates that about one-sixth of the assets under management by active portfolio managers in the U.S. large cap space is managed using quantitative strategies.

[2] Jegadeesh and Titman (1993) motivate their study of price momentum by noting that: "…a majority of the mutual funds examined by Grinblatt and Titman (1989; 1993) show a tendency to buy stocks that have increased in price over the previous quarter."

[3] The baseline 12-2 momentum strategy, described in more detail below, is based on a sort of individual stocks into decile portfolios based on the stocks' cumulative monthly returns from month $t-12$ through month $t-2$ (*i.e.*, skipping one month prior to portfolio formation). The strategy takes a long position in the value-weighted portfolio of the firms in the top decile, and a short position in value-weighted portfolio of the bottom decile stocks. The strategy is rebalanced monthly.

The pronounced leptokurtosis of the empirical distribution of momentum strategy returns suggests that these returns may be drawn from a mixture of distributions. We therefore develop a two state Hidden Markov Model (HMM) – a variant of the regime switching model of Hamilton (1989) – in which the hidden state can be "*calm*"or "*turbulent*." The states are persistent: successive months are more likely to be in the same state. Also, as we shall show momentarily, the joint distribution of the momentum and market returns differs significantly across the two states, facilitating our estimation of the HMM parameters.

The HMM model we propose is generally consistent with historical momentum crashes in our sample. When returns are generated by our hidden regime switching model, the probability of realizing a loss of more than 20% in 13 or more months in a sample of 978 months increases from 0.04% to 90%. The probability of incurring a loss of more than 79% increases from $9.95 \times 10^{-24}$ to 0.02%. This result, and others we detail below, suggests that our parsimonious 2-state model comes close to capturing the time variation in momentum strategy returns.[4]

The time variation in joint distribution of momentum and market returns has been discussed in the literature from a number of perspectives: Kothari and Shanken (1992) and Grundy and Martin (2001) observe that the market beta of momentum strategy returns is a function of past market returns. Rouwenhorst (1998) uncovers a nonlinearity in the momentum-market return relationship using methods suggested by Henriksson and Merton (1981). Boguth, Carlson, Fisher, and Simutin (2011) evaluate potential biases in estimating the momentum strategy alpha that may arise due to this nonlinearity. Daniel and Moskowitz (2011) observe that momentum strategies incur severe losses when the stock market recovers sharply following severe declines, and note the relationship between this nonlinearity and past market returns and volatility.

These empirical findings suggest that the sensitivity of momentum returns to stock market excess return – that is the beta of momentum returns – depends both on past realizations of market excess returns, and on the contemporaneous market excess return and its volatility. Typically, severe market declines and sharp recoveries occur during volatile market conditions, and high volatility periods tend to cluster. That suggests that the joint distribution of

[4]Based on one million simulations, in each of which 978 monthly returns are generated using estimated model parameters.

momentum and stock market returns depend on whether the market is volatile or calm. This leads us to use a HMM to describe the joint behavior of momentum and market returns. In our HMM specification, the joint distribution of momentum return and stock market return depends on the hidden state. This parsimonious HMM specification performs better than the other specifications we consider in predicting when large losses are more likely.

Most applications of the HMM in the empirical macro finance literature estimate changes in the hidden regimes using changes in the in expected values of the random variables. In contrast, we estimate the hidden state using second moments. Whereas reliable estimates of means require observations spanning a long period of time, variances and covariances can be estimated relatively precisely using more frequently sampled observations. We are therefore able to estimate the underlying parameters of the HMM relatively precisely even though our sample has only a few turbulent episodes each spanning relatively short time periods.

While we estimate the states using second moments, our estimation method uncovers differences in both the first moment and in third and fourth moments across the two states. Specifically, we find that in the turbulent state the momentum strategy returns have a lower mean, are more negatively skewed and exhibit increased kurtosis.

Finally, note that the quasi maximum likelihood estimation procedure we develop here provides consistent estimates of model parameters and their associated standard errors even though returns, by assumption, are non-normally distributed.[5]

While our use of the hidden Markov specification to capture sudden changes in the kurtosis of returns is novel, it has been widely used to model jumps and clustering of volatility in financial time series. For example, Calvet and Fisher (2008) find that their Markov-switching multifractal model (MSM), which is a generalization of the two state HMM we use, performs better than other models along several dimensions.[6] Our modeling approach allows for stochastic jumps in volatility, the importance of which has been emphasized by Todorov and Tauchen (2011), Andersen, Bondarenko, and Gonzalez-Perez (2011) and others.

The approach we take in identifying time periods when the hidden state is more likely to be turbulent is related to the approach of Bollerslev and Todorov (2011), who assume that

[5] See Appendix A.

[6] MSM is convenient for extending the number of volatility regimes with relatively few parameters. Since we only have a few turbulent episodes in our sample, the two state HMM that requires fewer parameters is preferable.

during time periods when more frequent medium sized jumps are more likely, less frequent large jumps are also more likely. Based on that assumption, for which they find support in the data, they are able to identify time periods with higher tail risk. Our approach is similar: We assume that tail risk as well as return volatilities and covariances jump in the unobserved turbulent state, and identify time periods when the hidden underlying state is more likely to be turbulent based on jumps in realized volatility. We find that the data support our assumptions.

Consistent with the extant findings in the literature, we find that when the hidden state is "*turbulent*", momentum strategy returns have both a strongly negative beta and the characteristics of a written call on the stock market index. These two features in particular allow us to accurately assess the hidden state.

Our estimation reveals that the distribution from which returns are drawn is more volatile when the market is in the unobserved "turbulent" state. We find that all the thirteen months with losses exceeding 20% occur during one of the 158 months when the probability of the market being in the turbulent state exceeds 70%. The probability of observing a loss of 20% or more in at least thirteen of 158 turbulent months is 94% and the large momentum strategy losses become less black swan-like.[7]

We see this graphically in Figure 1, which plots the *ex-ante* estimated probability that the hidden state is turbulent. Overlaid on this, we have placed a red dot for each of the 13 months in our sample in which the realized loss to the momentum strategy exceeded 20%. Notice that the estimated probability that the state is turbulent exceeds 70% during each of these 13 months.

As noted earlier, our HMM estimation procedure relies on the estimation of second moments, and in particular on the estimate of the covariance of the market return and the momentum strategy return. Daniel and Moskowitz (2011) observe that realized momentum strategy returns are considerably lower when the stock market recovers sharply following severe drawdowns (See Table 1). There are two reasons for this: first, firms which experience a high/low return at over a period when the realized market return is negative is more likely, from a Bayesian perspective, to be low/high beta. Second, following a prolonged bear

[7]These probabilities are based on one million simulations, in each of which 158 monthly returns are generated using the sample moments of the 158 months in our sample which have a predicted turbulent state probability exceeding 70%.

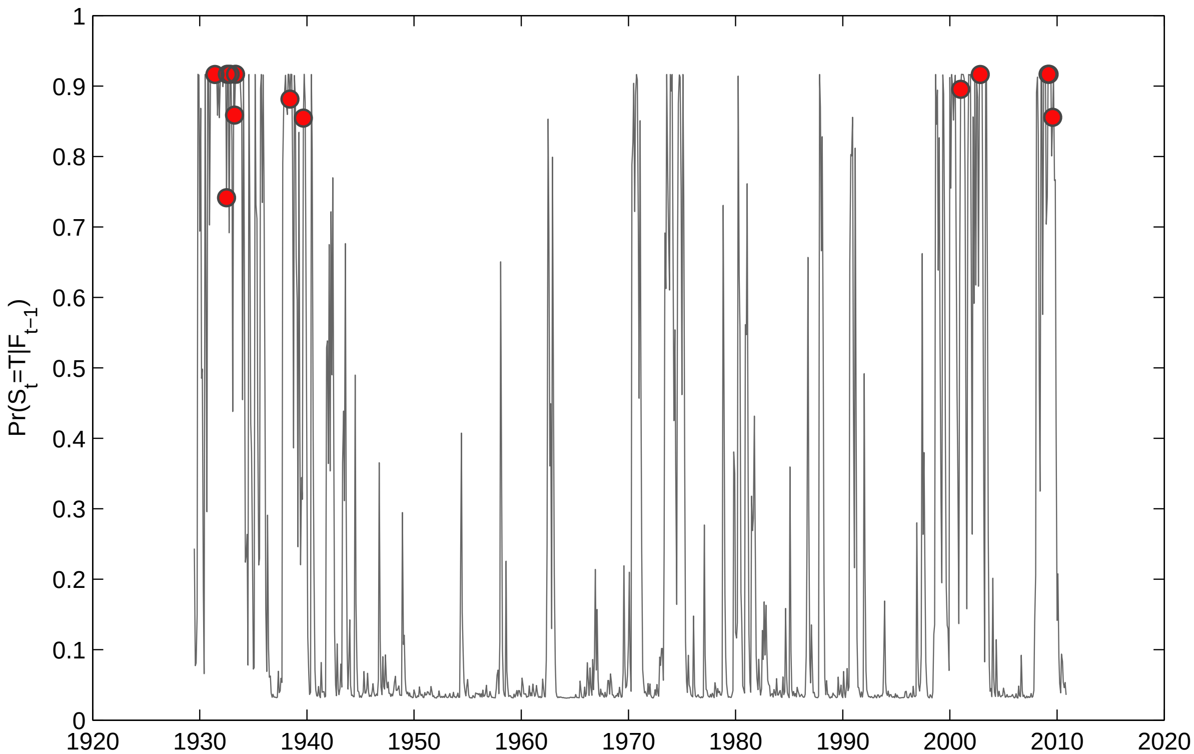


Figure 1: **Estimated Turbulent-State Probability and Momentum Crashes**

The solid line represents the HMM estimate of the *ex-ante* probability of hidden state being turbulent. The red dots indicate months in which the baseline momentum strategy lost of more than 20%.

market, the average firm's leverage increase, but the leverage of stocks in the lowest past returns decile increase even more, especially during turbulent times, with the result the beta of the momentum portfolio that shorts those stocks becomes significantly negative. Consequently, when the market recovers sharply, the momentum strategy incurs severe losses. Since momentum portfolios are formed by sorting stocks based on their returns over several months in the past, momentum portfolio betas become most negative (and hence most risky) some time after the hidden state becomes turbulent. Since the hidden states are persistent, our *ex-ante* estimate of the state probability also forecasts the future state.

We compare the performance of the HMM to other models. A GARCH model of time varying volatility is not as successful as the hidden regime switching models in identifying time periods when large losses are more likely to occur. Our findings are consistent with Ang and Timmerman (2011) who argue in favor of using regime switching models to capture abrupt changes in the statistical properties of financial market variables.

Table 1: NEGATIVE OUTLIERS

This table below lists the momentum "crash" months – defined as momentum strategy losses of more than 20%/month. $R_t^{\text{mom}}(\%)$ and $R_t^{\text{M}}(\%)$ are the realized momentum-strategy and market returns in the specified month. $\sum_{n=1}^{24} R_{t-n}^{\text{M}}(\%)$ is the cumulative market return over preceding 2 years, and *int.* is the time (in months) since the preceding crash month.

| MONTH | $R_t^{\text{mom}}(\%)$ | $R_t^{\text{M}}(\%)$ | $\sum_{n=1}^{24} R_{t-n}^{\text{M}}(\%)$ | *int.* |
|---|---|---|---|---|
| 1931/06 | -30.09 | 13.72 | -60.50 | |
| 1932/07 | -60.11 | 33.72 | -123.58 | 13 |
| 1932/08 | -78.96 | 36.75 | -93.63 | 1 |
| 1932/11 | -22.80 | -5.59 | -52.23 | 3 |
| 1933/04 | -42.33 | 38.27 | -63.68 | 5 |
| 1933/05 | -28.39 | 21.15 | -15.48 | 1 |
| 1938/06 | -33.14 | 23.61 | -24.49 | 61 |
| 1939/09 | -43.94 | 15.95 | -12.40 | 15 |
| 2001/01 | -42.10 | 3.41 | 3.28 | 736 |
| 2002/11 | -20.42 | 6.01 | -46.33 | 22 |
| 2009/03 | -39.32 | 8.76 | -60.92 | 76 |
| 2009/04 | -45.89 | 11.04 | -53.03 | 1 |
| 2009/08 | -24.80 | 3.18 | -28.89 | 4 |

We find that the monthly Sharpe Ratio of momentum returns in months with a predicted probability of being in the turbulent state exceeding 50% is -0.03. When those months are avoided the monthly Sharpe Ratio becomes 0.30, more than double the unconditional Sharpe Ratio. Momentum becomes more of an anomaly,[8] especially since the HMM based out of sample forecasts are about as good or better.

The rest of the paper is organized as follows. We provide a brief review of related literature in Section 2. We develop the econometric specifications of the HMM for characterizing momentum returns in Section 3. We discuss the empirical findings in Section 4 and examine alternative specifications in Section 5. We conclude in Section 6.

# 2 Related Literature

Levy (1967) was among the first academic articles to document the profitability of stock price

[8] As Hansen and Jagannathan (1991) observe, the high Sharpe Ratio portfolios pose a challenge to standard asset pricing models.

momentum strategies. However, Jensen (1967) raised several issues with the methodology employed by Levy (1967). Perhaps as a consequence, momentum received little attention in the academic literature until Jegadeesh and Titman (1993), whose long-short portfolio approach has proven rigorous and replicable. A number of studies have subsequently confirmed the Jegadeesh and Titman (1993) findings using data from markets in a number of countries (Rouwenhorst, 1998), and in a number of asset classes (Asness, Moskowitz, and Pedersen, 2008). As Fama and French (2008) observe, the "abnormal returns" associated with momentum are "pervasive".

Korajczyk and Sadka (2004) show that historical momentum profits remain positive after accounting for transaction costs. Chabot, Ghysels, and Jagannathan (2009) find that momentum strategies earned anomalous returns even during the Victorian era with very similar statistical properties, except for the January reversal, presumably because capital gains were not taxed in that period. Interestingly, they also find that momentum returns exhibited negative episodes once every 1.4 years with an average duration of 3.8 months per episode.

The additional literature on momentum strategies is vast, and can be grouped into three categories: (a) documentation of the momentum phenomenon across countries and asset classes (b) characterization of the statistical properties of momentum returns, both in the time-series and cross-section and (c) theoretical explanations for the momentum phenomenon. We make a contribution to (b). In what follows we provide review of only a few relevant articles that characterize the statistical properties of momentum returns, and refer the interested reader to the comprehensive survey of the momentum literature by Jegadeesh and Titman (2005).

Our work is directly related to several other papers in the literature. First, a number of authors, going back to Jegadeesh and Titman (1993), have noted that momentum strategies can experience severe losses over extended periods. In particular, Griffin, Ji, and Martin (2003) document that there are often periods of several months when momentum returns are negative.

In addition, a large number of papers have explored the time-varying nature of the risk of momentum strategies. Kothari and Shanken (1992) note that, for portfolios formed on the basis of past returns, the betas will be a function of past market returns. Using this intuition, Grundy and Martin (2001) show that the market beta of momentum strategy returns become negative when the stock market has performed poorly in the past. Rouwenhorst (1998) documents that momentum returns are nonlinearly related to contemporaneous market returns – specifically that the up market beta is less than their down market beta.[9] Daniel and Moskowitz (2011) find that this non-linearity is present only following market losses. Building on Cooper, Gutierrez, and Hameed (2004), who show that the expected returns associated with momentum profits depends on the state of the stock market, Daniel and Moskowitz (2011) also show that, after controlling for the dynamic risk of momentum strategies, average momentum returns are significantly lower following market losses and when measures of market volatility is high. They find evidence of these features not only in US equities, but also internationally, and in momentum strategies applied to commodity, currency, and country equity markets. Finally, they show that the factors combine to result in large momentum strategy losses in periods when the market recovers sharply following steep losses.[10]

# 3 Data and Econometric Specifications

## 3.1 Data

Price momentum strategies using stocks have been be constructed using variety of metrics. For this study we utilize the (12-2) momentum strategy decile portfolio returns available at Ken French's Data Library.[11] These portfolios are formed at the beginning of month $t$ by

[9]To our knowledge, Chan (1988) and DeBondt and Thaler (1987) first document that the market beta of a long-short winner-minus-loser portfolio is non-linearly related to the market return, though DeBondt and Thaler (1987) do their analysis on the returns of longer-term winners and losers as opposed to the shorter-term winners and losers we examine here. Rouwenhorst (1998) demonstrates the same non-linearity is present for long-short momentum portfolios in non-US markets. Finally Boguth, Carlson, Fisher, and Simutin (2011), building on the results of Jagannathan and Korajczyk (1986) and Glosten and Jagannathan (1994), note that the interpretation of the measures of abnormal performance in Chan (1988), Grundy and Martin (2001) and Rouwenhorst (1998) is problematic.

[10]Ambasta and Ben Dor (2010) observe a similar pattern for Barclays Alternative Replicator return that mimics the return on a broad hedge fund index. They find that when the hedge fund index recovers sharply from severe losses the replicator substantially underperforms the index. Also, Elavia and Kim (2011) emphasize the need for modeling changing risk for understanding the recent weak performance of quantitative equity investment managers who had over two decades of good performance.

[11]*http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html* The specific momentum decile portfolios we use are those designated by "10_Portfolios_Prior_12_2." Data for the CRSP value-

Table 2: Sample Moments of Momentum and Market Excess Returns
This table presents estimates of the first four moments of the monthly return distribution for 'mom', the zero-investment momentum portfolio used in the study, and for the CRSP value weighted stock portfolio net of the risk-free rate, $R^{\mathrm{M}}\,(=r^{\mathrm{M}}-r^{f})$. The sample period is 1929:07-2010:12.

| | Mean | Std.Dev | Skewness | Kurtosis |
|---|---|---|---|---|
| $R^{\mathrm{mom}}$ | 1.12% | 8.03% | -2.47 | 21.03 |
| $R^{\mathrm{M}}$ | 0.57% | 5.50% | 0.18 | 10.48 |

ranking each stock based its cumulative return over the 11 month period from month $t-12$ through month $t-2$.[12] The decile portfolio returns are the market-capitalization weighted portfolio of the stocks in that past return decile. Most of our analysis will concentrate on the zero-investment portfolio which is long the top past-return decile, and short the bottom decile. The long-short returns are defined as the difference between the top and bottom decile returns.

## 3.2 Characteristics of Momentum Returns

### 3.2.1 Descriptive Statistics

Table 2 presents estimates of the moments of the monthly momentum strategy returns, and of the CRSP value-weighted portfolio return, net of the risk-free rate. The monthly momentum strategy returns average 1.12% per month over this 978 month period, with a monthly Sharpe Ratio of 0.14. In contrast, the realized Sharpe ratio of the market over this period is only 0.10. Moreover, the alpha of momentum strategy is 1.70% per month with respect to the Fama and French (1993) three factor model. When the three Fama and French factors are combined with the momentum portfolio, the maximum achievable Sharpe Ratio rises to 0.28. Note that this is also considerably higher than the maximum Sharpe Ratio achievable with only three Fama French factors, 0.15.

While the momentum strategy has a higher in-sample Sharpe Ratio than the stock index

weighted market return and the risk-free rate is also taken from this data library.

[12] Skipping one month after the return measurement period is done both to be consistent with the momentum literature, and so as to minimize market microstructure effects and to avoid the short-horizon reversal effects documented in Jegadeesh (1990) and Lehmann (1990).

portfolio, Table 2 shows that it exhibits strong excessive kurtosis and negative skewness. The excess kurtosis is also evident from the infrequent but very large losses to the momentum strategy during the 978 months in the sample period. Given the low unconditional volatility of the momentum strategy, if returns were normal the probability of observing a month with a loss exceeding 42% in a sample of 978 months would be one in 25,000, and the probability of seeing five or more months with losses exceeding 42% would be almost zero. Yet the lowest five monthly returns in the sample are: -79%, -60%, -46%, -44%, and -42%, a rare black swan like occurrence from the perspective of someone who believes that the time series of monthly momentum returns are generated from an *i.i.d.* normal distribution.

When returns are simulated using our estimated HMM, the kurtosis is 8.64 for momentum return and 6.11 for market excess return. Simulated momentum returns exhibit some negative skewness, though not as much as in the sample data. Under the HMM model, large losses to momentum strategy become more likely: The probability of losses exceeding 42% in five or more months increases to 1.5% from almost zero; and the probability of losses exceeding 20% in thirteen or more months increases to 90% from 0.04%.[13]

In Section 3.3 we describe a hidden Markov model, based on the framework of Hamilton (1989), which we use to capture the behavior of momentum returns. Specifically, we model momentum returns with a mixture of normal distributions where the parameters of the normal distributions depend on the hidden state and contemporaneous market return in possibly nonlinear ways to accommodate the negative skewness of momentum strategy returns. Even though the distribution of returns – *conditional on the hidden state and the contemporaneous market return* – is normal, this model is able to capture the unconditional skewness and kurtosis of the momentum returns to some extent, and the probability of large losses become much more likely.

In what follows we first describe some of the salient characteristics of the joint distribution of momentum strategy and market returns documented in the literature. These features suggest that the beta of momentum strategy returns differs across turbulent and calm market conditions providing a rationale for hidden Markov model specification we use.

[13]Based on 1,000,000 simulations , in each of which 978 months returns are generated using estimated model parameters.

### 3.2.2 Momentum Beta and the Formation Period Market Return

Grundy and Martin (2001), building on the observation of Kothari and Shanken (1992), argue that the momentum portfolio beta will be a function of the market excess return over the measurement period.[14]

The intuition for the Kothari and Shanken (1992) result is as follows: if the market excess return is negative over the measurement period then, from a Bayesian perspective, the firms which earned the most negative return – *i.e.* the past losers which moved down with the market – are likely to have a higher beta than the past winners. Because the momentum portfolio is long past winners and short past losers, the intuition of Kothari and Shanken (1992) and Grundy and Martin (2001) suggests that the momentum portfolio beta should be negative following bear markets – when the market excess return was negative over the measurement period – and positive following bull markets.

We define the down market indicator $\mathrm{I}_t^D$ as having the value of one when the sum of the market excess returns $R_t^{\mathrm{M}}$ in the formation period was negative,[15] *i.e.* when:

$$\mathrm{I}_t^D \equiv \begin{cases} 1 & \text{if} \displaystyle\sum_{s=t-12}^{t-2} R_s^{\mathrm{M}} < 0 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

This suggests the following specification for the joint distribution of momentum and market excess returns:

$$R_t^{\mathrm{mom}} = \alpha + \left(\beta^0 + \beta^D \mathrm{I}_t^D\right) R_t^{\mathrm{M}} + \epsilon_t. \tag{2}$$

Estimation of this specification over our full sample yields:

| | $\alpha$ | $\beta^0$ | $\beta^D$ |
|---|---|---|---|
| estimate | 1.19 | 0.12 | **−1.27** |
| t-stat | 5.88 | 1.35 | **−8.37** |

[14] See pp. 195-198 of Kothari and Shanken (1992).

[15] We follow Grundy and Martin (2001) and use the arithmetic sum of the market excess returns during the formation period instead of the compounded total return during the formation period.

where the t-stats are computed using standard errors that allow for conditional heteroskedasticity. Consistent with the findings in the literature noted above, $\beta^D$ is economically and statistically significant, and abnormal return (intercept), controlling for the dynamic market risk, is still significantly positive.

### 3.2.3 Momentum Beta and Contemporaneous Market Returns

As pointed out in Section 2, a number of authors have observed the nonlinear nature of the relationship between the returns on past-return sorted portfolios and contemporaneous market returns.[16] In the language of Henriksson and Merton (1981), momentum's up-beta is different from its down-beta. We capture this option-like feature of the momentum portfolio with the following specification:

$$R_t^{\text{mom}} = \alpha + \left(\beta^0 + \beta^U \cdot \mathrm{I}_t^U\right) R_t^{\text{M}} + \epsilon_t, \tag{3}$$

where $\mathrm{I}_t^U$ is an indicator variable which is 1 when the contemporaneous market return is positive, and is zero otherwise:

$$\mathrm{I}_t^U \equiv \begin{cases} 1 & \text{if } \; R_t^M > 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Estimation of this specification over our full sample yields:

| | $\alpha$ | $\beta^0$ | $\beta^U$ |
|---|---|---|---|
| estimate | 3.27 | −0.07 | **−0.93** |
| t-stat | 6.29 | −0.64 | **−3.35** |

Notice that the estimated up-beta of our momentum portfolio is lower than the down-beta by 0.93, and is again strongly statistically significant. Also, here it is important to note that the estimated $\alpha$ is no longer a valid measure of abnormal performance, since the contemporaneous up-market indicator uses *ex-post* information.[17]

[16] See, in particular, footnote 9.

[17] See Jagannathan and Korajczyk (1986) and, more recently, Boguth, Carlson, Fisher, and Simutin (2011).

### 3.2.4 Dependence of Beta on Market Volatility

Boguth, Carlson, Fisher, and Simutin (2011) note the presence of a significant covariance between momentum's market beta and market volatility leads. Extending the approach of Jagannathan and Wang (1996), they calibrate the magnitude of the bias arising from this covariance.

To see whether volatility timing exists in our sample of momentum returns, we estimate the following regression equation:

$$R_t^{\text{mom}} = \alpha + \left(\beta^0 + \beta^U \cdot \text{I}_t^U + \beta^\sigma \cdot \widehat{\sigma}_t^{\text{M}}\right) R_t^{\text{M}} + \epsilon_t, \tag{5}$$

where $\text{I}_t^U$ is defined as in (4) and $\widehat{\sigma}_t^{\text{M}}$ is the sample standard deviation of daily returns for the value weighted portfolio during the month $t$. Estimation of this specification over our full sample yields:

| | $\alpha$ | $\beta^0$ | $\beta^U$ | $\beta^\sigma$ |
|---|---|---|---|---|
| estimate | 3.12 | 0.52 | **−1.07** | **−30.17** |
| t-stat | 7.57 | 2.72 | **−4.40** | **−3.36** |

Notice that an increase in the contemporaneous volatility of market excess returns significantly decreases the beta of momentum returns. As in the previous subsection, the estimated $\alpha$ is not a valid measure of abnormal performance, since the contemporaneous up-market indicator and the realized daily standard deviation uses *ex-post* information.

### 3.2.5 The State Dependence of Momentum's Optionality

Daniel and Moskowitz (2011) find that momentum portfolios incur large losses when the market recovers sharply following bear markets. For example, the worst five months with the largest losses on the momentum portfolio we discussed have the pattern in Table 3, consistent with the observations in Daniel and Moskowitz (2011).

Daniel and Moskowitz (2011) note that reason for this pattern is in part because the written call-features of momentum noted in the literature is particularly strong in bear markets. Daniel and Moskowitz (2011) therefore develop a specification in which beta of the momentum portfolio depends both on the long-term past return of the market, and the market's contemporaneous return. Consistent with this, we define another indicator variable.

Table 3: Patterns of Historical Momentum Crashes

| Month | $R_t^{\text{mom}}$ | $R_t^{\text{M}}$ | $R^{\text{M}}(t-24, t-1)$ |
|---|---|---|---|
| 1932/8 | -78.96 | 36.57 | -93.63 |
| 1932/7 | -60.11 | 33.72 | -123.58 |
| 2009/4 | -45.89 | 11.04 | -53.03 |
| 1939/9 | -43.94 | 15.95 | -12.40 |
| 1933/4 | -42.33 | 38.27 | -63.68 |

We define the bear-market indicator to be equal to 1 if the sum of the market excess returns over the preceding $L$ months is negative[18] :

$$\text{I}_t^B \equiv \begin{cases} 1 & \text{if } R^M(t-L, t-1) \equiv \sum_{s=t-L}^{t-1} R_s^{\text{M}} < 0 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

This motivates the following specification for the momentum return generating process.

$$R_t^{\text{mom}} = \alpha + \begin{pmatrix} \beta^0 \\ +\beta^B \cdot \text{I}_t^B \\ +\beta^R \cdot \text{I}_t^B \cdot \text{I}_t^U \end{pmatrix} R_t^{\text{M}} + \epsilon_t \tag{7}$$

Following Daniel and Moskowitz (2011), we set $L = 24$. Estimation of this specification yields:

| | $\alpha$ | $\beta^0$ | $\beta^B$ | $\beta^R$ |
|---|---|---|---|---|
| estimate | 1.71 | 0.19 | $-\mathbf{0.96}$ | $-\mathbf{0.80}$ |
| t-stat | 8.36 | 2.83 | $-\mathbf{7.52}$ | $-\mathbf{4.19}$ |

As can seen from the above table,

- $\beta^0 = 0.19$, i.e., the point estimate of the baseline beta is positive and statistically significant, though small.
- $\beta^B = -0.96$, i.e., the momentum beta is significantly negative following bear markets.
- $\beta^R = -0.80$, i.e., the momentum beta is much more negative when the contemporaneous market recovers following a bear market.

[18] As before, following Grundy and Martin (2001), we use the arithmetic sum of the market excess returns instead of the compounded total return.

#### 3.2.6 Summary

The sensitivity (beta) of momentum strategy return to stock market excess return depends on: (1) the return on the market during the momentum portfolio formation period; (2) whether the market has declined severely during the past two years; (3) the volatility of the market; and (4) the contemporaneous return on the market. One source of these patterns is the high leverage of past losers, particularly following large market declines. This high leverage would mean that the returns of past losers would both have a high beta and exhibit pronounced call option-like features during times when the stock market is depressed. The momentum portfolio, which is short these past losers, will behave as though it is short an out-of-the money call: it will have a large negative beta, which is even more negative when the market recovers sharply, thus resulting in a momentum crash. Since the market is often turbulent during such time periods, we assume that there is a single hidden state variable that can capture all these effects, providing a rationale for the hidden Markov model of momentum returns we describe in the next section.

### 3.3 A Hidden Markov Model of Momentum Strategy Returns

Let $S_t$ denote the unobserved random underlying state of the economy, and $s_t$ the realized state of the economy. We assume that there are two possible states, one is Calm (C) and the other Turbulent (T). Our specification for the momentum portfolio return generating process is as follows:

$$R_t^{\text{mom}} = \alpha(S_t) + \left( \beta^0(S_t) + \beta^U(S_t) \cdot \text{I}_t^U \right) R_t^{\text{M}} + \sigma_{\text{mom}}(S_t)\epsilon_t^{\text{mom}}. \tag{8}$$

Here $\epsilon_t^{\text{mom}}$ denotes the standardized momentum strategy residual return – a sequence of independent random variables with zero mean and unit variance, and $\sigma_{\text{mom}}(S_t)$ denotes the standard deviation of the momentum strategy residual return. Note that the volatility of the residual return is a function of the hidden state. The specification in (8) is the same as in (3) except for the dependence of model parameters on the hidden state. We do not explicitly model the dependence of the momentum beta on $\text{I}_t^D$, $\widehat{\sigma}_t^{\text{M}}$ or $\text{I}_t^B$ as in the specifications in (2), (5) and (7). Since we expect the hidden state to sufficiently summarize the relevant market conditions, we let the model parameters be a function of only the hidden state.[19] In addition, we allow momentum beta to depend on contemporaneous market excess return in nonlinear ways, i.e., be a function of the up market indicator variable, $\mathrm{I}_t^U$ defined in (4), in order to capture the option-like features of momentum returns.

We further specify that the mean and the variance of the market excess return are also functions of the hidden state, specifically

$$R_t^{\mathrm{M}} = \mu\left(S_t\right) + \sigma_{\mathrm{M}}\left(S_t\right)\epsilon_t^{\mathrm{M}}, \tag{9}$$

where $\epsilon_t^{\mathrm{M}}$ denotes the standardized unanticipated market returns – a sequence of independent random variables with zero mean and unit variance – and $\sigma_{\mathrm{M}}(S_t)$ denotes the standard deviation of the unanticipated market return.

## 3.4 Maximum Likelihood Estimation

We estimate the parameters of the hidden Markov model by assuming that the standardized momentum strategy residual return and the unanticipated market return are jointly normally distributed, even though they are not, by using the maximum likelihood method described in Hamilton (1989). This approach is equivalent to using the GMM to estimate the model parameters based on the information in the conditional first and second moments alone. In Appendix A, we show that the consistency and the asymptotic normality of the estimates obtained via maximum likelihood does not require that $\epsilon_t^{\mathrm{M}}$ and $\epsilon^{\mathrm{mom}}$ be jointly normally distributed.[20]

We begin by defining the vector of observable variables of interest at time $t$, $Y_t$, as:

$$Y_t = \left(\mathrm{I}_t^U, R_t^{\mathrm{mom}}, R_t^{\mathrm{M}}\right)',$$

We let $y_t$ denote the realized value of $Y_t$. $S_t$ denotes the unobservable random state at time $t$ which, in our setting, is either *calm* or *turbulent*. $s_t$ denotes the particular realized value

[19] As we will see later, letting beta depend on past market returns and volatilities does not improve the model's performance.

[20] For consistency and asymptotic normality of the estimates, what is required is that equations (8) and (9) continue to hold with the assumption that a sequence of random vectors, $\left[\epsilon_t^{\mathrm{M}}\ \epsilon_t^{\mathrm{mom}}\right]'$, satisfy three minimal conditions: zero mean, unit variance, and independence across time.

of the state at date $t$. $\Pr(S_t = s_t | S_{t-1} = s_{t-1})$ denotes the transition probability of moving from state $s_{t-1}$ at time $t-1$ to state $s_t$ at time $t$.[21] Finally, $\mathcal{F}_{t-1}$ denotes the information set at time $t-1$, *i.e.*, $\{y_{t-1}, \cdots, y_1\}$.

The evolution of the two hidden states are determined by the transition probabilities from one state to another. In our estimation, we set $\Pr(S_0 = s_0)$ to the unconditional probabilities, corresponding to the transition probabilities.

Suppose we know the value of $\Pr(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1})$. Then, $\Pr(S_t = s_t | \mathcal{F}_{t-1})$ is given by:

$$\begin{aligned} \Pr(S_t = s_t | \mathcal{F}_{t-1}) &= \sum_{S_{t-1}} \Pr(S_t = s_t, S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}) \\ &= \sum_{S_{t-1}} \Pr(S_t = s_t | S_{t-1} = s_{t-1}, \mathcal{F}_{t-1}) \Pr(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}) \\ &= \sum_{S_{t-1}} \Pr(S_t = s_t | S_{t-1} = s_{t-1}) \Pr(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}), \end{aligned} \tag{10}$$

The third equality holds since the transition probabilities depend only on the hidden state. We can compute the expression on the right side of equation (10) using the elements of the transition probability matrix, $\Pr(S_t = s_t | S_{t-1} = s_{t-1})$ and $\Pr(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1})$.

Next, we can compute the joint conditional distribution of the hidden state at time $t$, $s_t$ and observable variables at time $t$, $y_t$ as follows.

$$\Pr(y_t, S_t = s_t | \mathcal{F}_{t-1}) = \Pr(y_t | S_t = s_t, \mathcal{F}_{t-1}) \Pr(S_t = s_t | \mathcal{F}_{t-1}) \tag{11}$$

The second term of this expression, $\Pr(S_t = s_t | \mathcal{F}_{t-1})$, is given by (10). The first term is the state dependent likelihood of $y_t$ which, under the distributional assumptions from (8) and (9), is given by

$$\begin{aligned} \Pr(y_t | S_t = s_t, \mathcal{F}_{t-1}) = &\frac{1}{\sigma_{\text{mom}}(s_t)\sqrt{2\pi}} \exp\left\{-\frac{(\epsilon_t^{\text{mom}})^2}{2}\right\} \\ &\times \frac{1}{\sigma_{\text{M}}(s_t)\sqrt{2\pi}} \exp\left\{-\frac{(\epsilon_t^{\text{M}})^2}{2}\right\} \end{aligned} \tag{12}$$

[21] Here, we use $\Pr(x)$ to denote the probability of the event $x$ when $x$ is discrete, and the probability density of $x$ when $x$ is continuous.

where

$$\epsilon_t^{\text{mom}}(s_t) = \frac{1}{\sigma_{\text{mom}}(s_t)} \left( R_t^{\text{mom}} - \alpha(s_t) - \beta\left(S_t, \mathrm{I}_t^U\right) R_t^{\text{M}} \right) \tag{13}$$

$$\epsilon_t^{\text{M}}(s_t) = \frac{1}{\sigma_{\text{M}}(s_t)} \left( R_t^{\text{M}} - \mu(s_t) \right). \tag{14}$$

The likelihood of $y_t$ given $\mathcal{F}_{t-1}$ is:

$$\Pr(y_t|\mathcal{F}_{t-1}) = \sum_{S_t} \Pr(y_t, S_t = s_t|\mathcal{F}_{t-1}). \tag{15}$$

where the joint likelihood is be calculated using equation (11).

Finally Bayes' rule gives the state probability at $t$ as a function of the contemporaneous information set:

$$\begin{aligned} \Pr(S_t = s_t|\mathcal{F}_t) &= \Pr(S_t = s_t|y_t, \mathcal{F}_{t-1}) \\ &= \frac{\Pr(y_t, S_t = s_t|\mathcal{F}_{t-1})}{\Pr(y_t|\mathcal{F}_{t-1})} \end{aligned} \tag{16}$$

Using (16) to compute $\Pr(S_t = s_t|\mathcal{F}_t)$ and the algorithm described above, we can compute $\Pr(y_{t+1}|\mathcal{F}_t)$. In this way, we can generate the likelihood $\Pr(y_{t+1}|\mathcal{F}_t)$ for $t = 0, 1, \cdots, T-1$ and compute the log-likelihood $\mathcal{L} = \sum_{t=1}^{T} \log(\Pr(y_t|\mathcal{F}_{t-1}))$ given specific values for the unknown parameters that we need to estimate. Now consider the log likelihood function of the sample given by:

$$\mathcal{L} = \sum_{t=1}^{T} \log(\Pr(y_t|\mathcal{F}_{t-1})), \tag{17}$$

which may be maximized numerically to form estimates of the parameter $\theta \in \Theta$, where $\Theta$ is a compact set which contains the true parameter of $\theta^0$,

$$\theta^0 = \left\{ \begin{array}{c} \alpha(C), \beta^0(C), \beta^U(C), \sigma_{\text{mom}}(C) \\ \alpha(T), \beta^0(T), \beta^U(T), \sigma_{\text{mom}}(T) \\ \mu(C), \sigma_{\text{M}}(C), \mu(T), \sigma_{\text{M}}(T) \\ \Pr(S_t = C|S_t = C), \Pr(S_t = T|S_t = T) \end{array} \right\}.$$

We estimate the model parameters by maximizing the log likelihood given by (17), *i.e.*,

$$\theta_{ML} = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta), \tag{18}$$

and use quasi maximum likelihood standard errors for inference.

# 4 Empirical Results

Table 4 summarizes the maximum likelihood estimates of the model parameters. The baseline beta $\beta^0$ is positive for the unobserved *calm* state and negative for the unobserved *turbulent* state. $\beta^U$, which captures the written-call like characteristic of the momentum portfolio, is present in both states but much stronger in the *turbulent* state. The momentum return beta during up-markets is -0.11($= \beta^0_C + \beta^U_C = 0.41 - 0.52 = -0.11$) in the *calm* state, significantly smaller in magnitude than -1.54($= \beta^0_T + \beta^U_T = -0.26 - 1.28 = -1.54$)), the corresponding value of beta in the *turbulent* state. While both the calm and turbulent states are persistent, the probability of remaining in the calm state starting from the calm state is higher.

Table 4: Maximum Likelihood Estimates of HMM Parameters

This table presents the parameters of the HMM model, estimated using the momentum strategy and the market excess returns over the full sample. QML standard errors are used to compute t-stats. The ML estimates of $\alpha$, $\sigma_{\text{mom}}$ and $\sigma_{\text{M}}$ are each multiplied by 100. The maximized value of the Log-Likelihood is $3.01 \times 10^3$; and the value of BIC is $-5.93 \times 10^3$.

| | Hidden State | | | |
|---|---|---|---|---|
| | Calm | | Turbulent | |
| Parameter | MLE | t-stat | MLE | t-stat |
| $\alpha$ ($\times 10^2$) | 2.04 | 6.98 | 4.02 | 3.37 |
| $\beta^0$ | 0.41 | 2.59 | $-0.26$ | $-1.63$ |
| $\beta^U$ | $-0.52$ | $-2.46$ | $-1.28$ | $-4.31$ |
| $\sigma_{\text{mom}}$ ($\times 10^2$) | 4.32 | 15.49 | 11.05 | 11.62 |
| $\mu$ | 0.98 | 6.24 | $-0.70$ | $-1.14$ |
| $\sigma_{\text{M}}$ ($\times 10^2$) | 3.62 | 24.62 | 9.04 | 9.42 |
| $\Pr(S_t = s_{t-1} \mid S_{t-1} = s_{t-1})$ | 0.97 | 57.87 | 0.92 | 15.39 |

Table 5 provides the t-stats for the difference in parameter values across the *calm* and the *turbulent* states. $\beta^0$, the baseline beta, is much higher in the *calm* state. $\beta^U$ – the parameter that captures the written call feature of the momentum strategy return is accentuated in the *turbulent* state, resulting in a large negative momentum strategy beta in up-markets. The volatility of the market excess return as well as the residual return of the momentum portfolio are significantly different across two hidden states. They are more than twice as volatile in the turbulent state when compared to the calm state. These differences in parameter values across the two hidden states help us infer the hidden state based on past observations on momentum and market returns. The hidden state being persistent helps in forecasting which state is more likely to prevail in the immediate future.

Table 5: Differences in Parameter Values Across Hidden States

| | MLE | t-stat |
|---|---|---|
| $\alpha(C) - \alpha(T)$ | −1.99 | −1.60 |
| $\beta^0(C) - \beta^0(T)$ | 0.67 | 3.07 |
| $\beta^U(C) - \beta^U(T)$ | 0.76 | 2.01 |
| $\sigma_{\text{mom}}(C) - \sigma_{\text{mom}}(T)$ | −6.73 | −6.95 |
| $\mu(C) - \mu(T)$ | 1.68 | 2.55 |
| $\sigma_{\text{M}}(C) - \sigma_{\text{M}}(T)$ | −5.42 | −5.97 |

Table 6 gives the number of positive and negative momentum strategy returns exceeding a threshold level when the predicted probability of the next month being in the turbulent state is $p$ or more, where $p$ takes any of the values from 10% to 90% in steps of 10%, for various threshold levels. Each entry represents the fraction of the months with large gains(losses) exceeding a given threshold level during months when the underlying state is *turbulent* with a probability exceeding $p$.

It is natural to classify months as being turbulent when the unobserved underlying state is *turbulent* with a probability exceeding 50%. When we estimate the model using the entire sample, all 13 months in which the momentum strategy losses exceed 20% occur during months in which $\Pr(S_t = T|\mathcal{F}_{t-1}) > 50\%$.[22] However, there are 11 months in which the momentum strategy return exceeds 20%, and only 8 of those 11 months have $\Pr(S_t = T|\mathcal{F}_{t-1}) > 50\%$. That is, the predicted probability of the next month being in the turbulent state is more informative about the likelihood of large losses than large gains.

[22]There are a total of 199 months, out of a total of 978 months in the sample, in which $\Pr(S_t = T|\mathcal{F}_{t-1}) >$ 50%.

Table 6: LARGE MOMENTUM STRATEGY LOSSES/GAINS IN TURBULENT/CALM MONTHS – IN SAMPLE

| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ is more than | # Losses Captured/Total #Losses $\leq -10\%$ | $\leq -12.5\%$ | $\leq -15\%$ | $\leq -17.5\%$ | $\leq -20\%$ | # of Months |
|---|---|---|---|---|---|---|
| 90% | 19/53 | 16/35 | 14/31 | 10/21 | 7/13 | 87 |
| 80% | 27/53 | 24/35 | 21/31 | 17/21 | 12/13 | 131 |
| 70% | 30/53 | 26/35 | 23/31 | 18/21 | 13/13 | 158 |
| 60% | 31/53 | 27/35 | 24/31 | 19/21 | 13/13 | 188 |
| 50% | 32/53 | 27/35 | 24/31 | 19/21 | 13/13 | 199 |
| 40% | 34/53 | 28/35 | 25/31 | 20/21 | 13/13 | 225 |
| 30% | 37/53 | 29/35 | 26/31 | 20/21 | 13/13 | 252 |
| 20% | 40/53 | 31/35 | 28/31 | 20/21 | 13/13 | 291 |
| 10% | 43/53 | 33/35 | 30/31 | 21/21 | 13/13 | 342 |

| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ is more than | # Gains captured/Total #Gains $\geq 10\%$ | $\geq 12.5\%$ | $\geq 15\%$ | $\geq 17.5\%$ | $\geq 20\%$ | # of Months |
|---|---|---|---|---|---|---|
| 90% | 21/66 | 12/41 | 6/26 | 4/14 | 3/11 | 87 |
| 80% | 29/66 | 18/41 | 12/26 | 7/14 | 5/11 | 131 |
| 70% | 33/66 | 22/41 | 16/26 | 10/14 | 8/11 | 158 |
| 60% | 39/66 | 27/41 | 18/26 | 10/14 | 8/11 | 188 |
| 50% | 40/66 | 28/41 | 18/26 | 10/14 | 8/11 | 199 |
| 40% | 40/66 | 28/41 | 18/26 | 10/14 | 8/11 | 225 |
| 30% | 44/66 | 29/41 | 19/26 | 11/14 | 9/11 | 252 |
| 20% | 47/66 | 31/41 | 20/26 | 11/14 | 9/11 | 291 |
| 10% | 50/66 | 31/41 | 20/26 | 11/14 | 9/11 | 342 |

Table 7 gives the out of sample results when we predict the probability of hidden state being turbulent in month $t+2$ based on information available at $t$, i.e., $\Pr(S_t = T \mid \mathcal{F}_{t-2})$. As can be seen, all the 13 months with losses exceeding 20% occur in one of the 252 months in which the predicted probability of being in the turbulent state exceeds 30%, and 11 of the 13 occur in months when the probability exceeds 60%. That is, the forecastability associated with the HMM model is persistent: we are able to predict the likelihood of extreme losses two months ahead with some accuracy.

Next, we examine the out of sample performance of the hidden Markov model to identify months when large losses to the momentum strategy are more likely based on *ex-ante* infor-

Table 7: Large Momentum Strategy Losses/Gains in Turbulent/Calm Months – In Sample(One Month Lag)

| $\Pr(S_t = T \mid \mathcal{F}_{t-2})$ | # Losses Captured/Total #Losses | | | | | # |
|---|---|---|---|---|---|---|
| is more than | $\leq -10\%$ | $\leq -12.5\%$ | $\leq -15\%$ | $\leq -17.5\%$ | $\leq -20\%$ | of Months |
| 90% | 0/53 | 0/35 | 0/31 | 0/21 | 0/13 | 0 |
| 80% | 15/53 | 14/35 | 13/31 | 10/21 | 6/13 | 106 |
| 70% | 20/53 | 18/35 | 16/31 | 12/21 | 8/13 | 143 |
| 60% | 28/53 | 25/35 | 23/31 | 18/21 | 11/13 | 177 |
| 50% | 32/53 | 27/35 | 25/31 | 19/21 | 12/13 | 197 |
| 40% | 34/53 | 27/35 | 25/31 | 19/21 | 12/13 | 223 |
| 30% | 37/53 | 29/35 | 27/31 | 21/21 | 13/13 | 252 |
| 20% | 40/53 | 31/35 | 29/31 | 21/21 | 13/13 | 293 |
| 10% | 43/53 | 34/35 | 31/31 | 21/21 | 13/13 | 370 |

| $\Pr(S_t = T \mid \mathcal{F}_{t-2})$ | # Gains captured/Total #Gains | | | | | # |
|---|---|---|---|---|---|---|
| is more than | $\geq 10\%$ | $\geq 12.5\%$ | $\geq 15\%$ | $\geq 17.5\%$ | $\geq 20\%$ | of Months |
| 90% | 0/66 | 0/41 | 0/26 | 0/14 | 0/11 | 0 |
| 80% | 23/66 | 18/41 | 11/26 | 6/14 | 5/11 | 106 |
| 70% | 28/66 | 21/41 | 13/26 | 8/14 | 6/11 | 143 |
| 60% | 33/66 | 24/41 | 15/26 | 9/14 | 7/11 | 177 |
| 50% | 35/66 | 25/41 | 15/26 | 9/14 | 7/11 | 197 |
| 40% | 37/66 | 26/41 | 16/26 | 10/14 | 8/11 | 223 |
| 30% | 40/66 | 26/41 | 16/26 | 10/14 | 8/11 | 252 |
| 20% | 44/66 | 29/41 | 19/26 | 12/14 | 9/11 | 293 |
| 10% | 49/66 | 32/41 | 22/26 | 13/14 | 10/11 | 370 |

mation in real time. For that purpose we use an expanding window to estimate the model parameters and that gives us real time predicted probability of the next month being in the turbulent state for 400 months. In particular, for each month $t = T - 399, \cdots, T$, we estimate the model parameters using maximum likelihood using data for the months $\{1, \cdots, t-1\}$. The first out-of-sample month is September 1977 and the last is December 2010.

From Table 8 we can see that, over the out-of-sample period, there were 5 months in which the momentum strategy lost more than 20%. All of these large losses occurred during the 79 months when the probability of the market being in the turbulent state exceeds 50%. Indeed, *all of these losses occurred when the turbulent state probability exceeds 90%.* In

Table 8: Large Momentum Strategy Losses/Gains in Turbulent/Calm Months – Out of Sample

| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ is more than | # Losses captured/Total #Losses | | | | | # of Months |
|---|---|---|---|---|---|---|
| | $\leq -10\%$ | $\leq -12.5\%$ | $\leq -15\%$ | $\leq -17.5\%$ | $\leq -20\%$ | |
| 90% | 12/26 | 10/16 | 8/12 | 6/7 | 5/5 | 41 |
| 80% | 13/26 | 11/16 | 8/12 | 6/7 | 5/5 | 53 |
| 70% | 13/26 | 11/16 | 8/12 | 6/7 | 5/5 | 59 |
| 60% | 14/26 | 11/16 | 8/12 | 6/7 | 5/5 | 72 |
| 50% | 14/26 | 11/16 | 8/12 | 6/7 | 5/5 | 79 |
| 40% | 14/26 | 11/16 | 8/12 | 6/7 | 5/5 | 86 |
| 30% | 14/26 | 11/16 | 8/12 | 6/7 | 5/5 | 91 |
| 20% | 15/26 | 12/16 | 9/12 | 6/7 | 5/5 | 107 |
| 10% | 16/26 | 13/16 | 10/12 | 7/7 | 5/5 | 132 |

| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ is more than | # Gains captured/Total #Gains | | | | | # of Months |
|---|---|---|---|---|---|---|
| | $\geq 10\%$ | $\geq 12.5\%$ | $\geq 15\%$ | $\geq 17.5\%$ | $\geq 20\%$ | |
| 90% | 7/29 | 5/18 | 3/12 | 1/5 | 0/4 | 41 |
| 80% | 8/29 | 6/18 | 4/12 | 2/5 | 1/4 | 53 |
| 70% | 8/29 | 6/18 | 4/12 | 2/5 | 1/4 | 59 |
| 60% | 12/29 | 9/18 | 6/12 | 3/5 | 2/4 | 72 |
| 50% | 15/29 | 11/18 | 6/12 | 3/5 | 2/4 | 79 |
| 40% | 16/29 | 12/18 | 7/12 | 3/5 | 2/4 | 86 |
| 30% | 17/29 | 13/18 | 8/12 | 4/5 | 3/4 | 91 |
| 20% | 18/29 | 13/18 | 8/12 | 4/5 | 3/4 | 107 |
| 10% | 21/29 | 14/18 | 9/12 | 4/5 | 3/4 | 132 |

contrast, it is difficult to predict when large gains are more likely: There were 4 months in which gains exceeded 20%, but only 2 of them occurred during months in which the probability of being in the turbulent state exceeded 50%. Graphically, this can been seem in both Figure 2 and Figure 3: the large losses to the momentum strategy all occur when the turbulent state probability is high. However, gains exceeding 20% are not so strongly associated with the turbulent state probability.

This asymmetry is probably due to the fact that momentum crashes tend to occur when the market recovers from steep losses but there are no corresponding gains when market continues to depreciate instead of recovering.

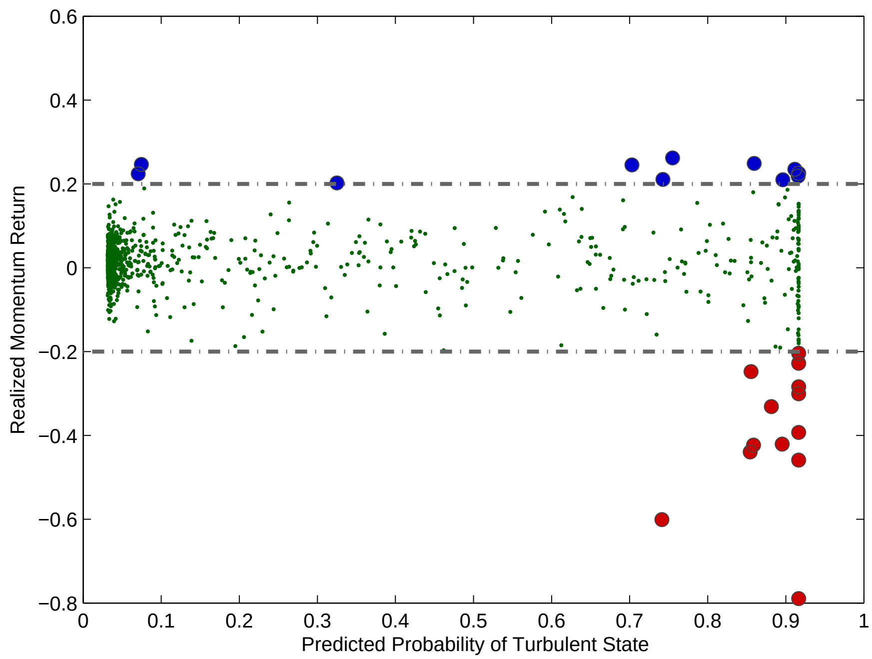


Figure 2: **Momentum Returns & Turbulent-State Probabilities (in-sample)**

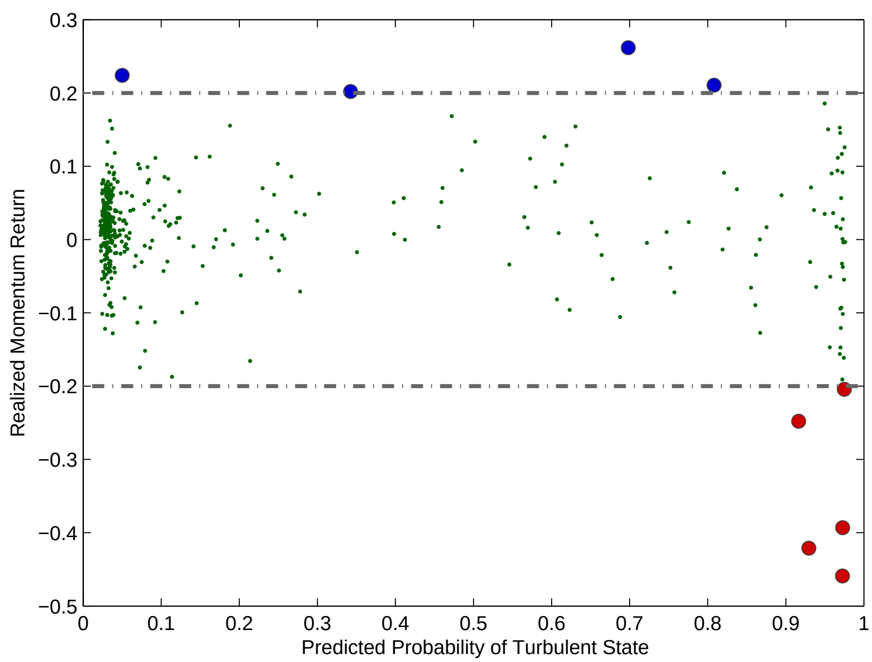


Figure 3: **Momentum Returns & Turbulent-State Probabilities (out-of-sample)**

Each point in the above figures represents a momentum strategy return. The location along the x-axis represents the estimated probability that the state is turbulent at the beginning of the month. The location of the point along the y-axis represents the momentum strategy return in the following month. Blue points represent realized returns $> 20\%$, and red points represent realized returns of $< 20\%$. The probabilities in Figure 2 are based on in-sample estimates, while the probabilities in Figure 3 are based on out-of-sample estimates.

Panel A of Table 9 shows that the properties of momentum returns depend on whether the hidden state is more likely to be turbulent or calm. Here, for each row of the Table, we classify each month of the sample based on whether, and the beginning of each month, the estimated turbulent state probability is above or below the stated threshold probability.

The Sharpe Ratio of momentum returns in months when the hidden state is more likely to be calm is more than double of the Sharpe Ratio for all months in the sample. For example, with a threshold probability of 50%, 199 months are classified turbulent and 779 months are calm. Momentum returns are on average negative though are not statistically significantly so during turbulent months, and almost three times as volatile as calm months. For comparison, we also report the corresponding statistics for the market excess return in the panel B of table 9. The average returns for calm and turbulent months are not as much different for the market, which is consistent with the observations in Breen, Glosten, and Jagannathan (1989).

The properties of momentum and market returns during the out of sample period is summarized in Table 10. With the threshold level of 50% for the probability for the month being in the turbulent state, about 80% of the sample months are classified as *calm* and 20% of the months are classified as *turbulent*. The characteristics of momentum returns for the out-of-sample months are similar to that during the in-sample results. As we increase threshold level for the probability of the underlying state being turbulent in a given month, the sample mean of the momentum returns during turbulent months decrease while the sample mean of the market excess returns increase. With the threshold level of 90% for the probability for being in the turbulent state, which classifies 10% of the 400 out-of-sample as *turbulent*, the Sharpe ratio of momentum returns for turbulent months becomes -0.26. The Sharpe Ratio of market returns during those months is 0.38, rather high. This is to be expected since momentum crashes tend to occur when the market recovers, as observed by Daniel and Moskowitz (2011).

In Table 11 we report the properties of momentum returns where we classify month when the hidden state is turbulent with a probability greater than 50%. When compared to the properties of momentum returns for the entire sample returns are less leptokurtic, and the kurtosis for the turbulent months are more than that for calm months. The higher kurtosis relative to normal is to be expected since we do not know what the true underlying state is;

Table 9: Returns in Turbulent and Calm Months

Months with $\Pr(S_t = T | \mathcal{F}_{t-1})$ exceeding the threshold level are classified as *turbulent*, and all other months are classified as *calm*.

Panel A: Momentum Returns

| | Forecasted State | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Calm | | | | Turbulent | | | |
| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ | Mean | SD | SR | Months | Mean | SD | SR | Months |
| 10% | 1.64 | 4.62 | 0.35 | 636 | 0.16 | 11.99 | 0.01 | 342 |
| 20% | 1.62 | 4.82 | 0.34 | 687 | -0.06 | 12.67 | -0.00 | 291 |
| 30% | 1.58 | 4.93 | 0.32 | 726 | -0.20 | 13.36 | -0.01 | 252 |
| 40% | 1.58 | 5.05 | 0.31 | 753 | -0.43 | 13.89 | -0.03 | 225 |
| 50% | 1.53 | 5.14 | 0.30 | 779 | -0.47 | 14.54 | -0.03 | 199 |
| 60% | 1.54 | 5.17 | 0.30 | 790 | -0.62 | 14.86 | -0.04 | 188 |
| 70% | 1.59 | 5.32 | 0.30 | 820 | -1.30 | 15.72 | -0.08 | 158 |
| 80% | 1.54 | 5.91 | 0.26 | 847 | -1.60 | 15.79 | -0.10 | 131 |
| 90% | 1.35 | 6.79 | 0.20 | 891 | -1.27 | 15.79 | -0.08 | 87 |

Panel B: Market Excess Returns

| | Forecasted State | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Calm | | | | Turbulent | | | |
| $\Pr(S_t = T \mid \mathcal{F}_{t-1})$ | Mean | SD | SR | Months | Mean | SD | SR | Months |
| 10% | 0.68 | 4.10 | 0.17 | 636 | 0.37 | 7.45 | 0.05 | 342 |
| 20% | 0.61 | 4.25 | 0.14 | 687 | 0.49 | 7.71 | 0.06 | 291 |
| 30% | 0.56 | 4.31 | 0.13 | 726 | 0.60 | 8.02 | 0.08 | 252 |
| 40% | 0.55 | 4.32 | 0.13 | 753 | 0.64 | 8.33 | 0.08 | 225 |
| 50% | 0.61 | 4.39 | 0.14 | 779 | 0.44 | 8.60 | 0.05 | 199 |
| 60% | 0.55 | 4.40 | 0.13 | 790 | 0.65 | 8.75 | 0.07 | 188 |
| 70% | 0.50 | 4.45 | 0.11 | 820 | 0.95 | 9.22 | 0.10 | 158 |
| 80% | 0.52 | 4.63 | 0.11 | 847 | 0.94 | 9.38 | 0.10 | 131 |
| 90% | 0.56 | 5.07 | 0.11 | 891 | 0.65 | 8.82 | 0.07 | 87 |

Table 10: Returns in Turbulent and Calm Months (Out of Sample)
Months with $\Pr(S_t = T | \mathcal{F}_{t-1})$ exceeding the threshold level are counted as *turbulent* and other months are counted as *calm.*

Panel A: Momentum Returns

| | Forecasted State | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Calm | | | | Turbulent | | | |
| $\Pr(S_t = T \| \mathcal{F}_{t-1})$ | Mean | SD | SR | Months | Mean | SD | SR | Months |
| 10% | 1.63 | 5.10 | 0.32 | 268 | 0.40 | 11.14 | 0.04 | 132 |
| 20% | 1.59 | 5.30 | 0.30 | 293 | 0.20 | 11.88 | 0.02 | 107 |
| 30% | 1.54 | 5.37 | 0.29 | 309 | 0.15 | 12.60 | 0.01 | 91 |
| 40% | 1.61 | 5.45 | 0.30 | 314 | -0.20 | 12.74 | -0.02 | 86 |
| 50% | 1.72 | 5.49 | 0.31 | 321 | -0.80 | 13.05 | -0.06 | 79 |
| 60% | 1.83 | 5.55 | 0.33 | 328 | -1.53 | 13.31 | -0.11 | 72 |
| 70% | 1.87 | 5.82 | 0.32 | 341 | -2.55 | 13.65 | -0.19 | 59 |
| 80% | 1.84 | 5.81 | 0.32 | 347 | -2.84 | 14.29 | -0.20 | 53 |
| 90% | 1.82 | 5.93 | 0.31 | 359 | -4.03 | 15.39 | -0.26 | 41 |

Panel B: Market Excess Returns

| | Forecasted State | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Calm | | | | Turbulent | | | |
| $\Pr(S_t = T \| \mathcal{F}_{t-1})$ | Mean | SD | SR | Months | Mean | SD | SR | Months |
| 10% | 0.70 | 6.79 | 0.10 | 268 | 0.74 | 6.42 | 0.12 | 132 |
| 20% | 0.61 | 6.90 | 0.09 | 293 | 0.99 | 6.00 | 0.16 | 107 |
| 30% | 0.71 | 7.10 | 0.10 | 309 | 0.73 | 4.96 | 0.15 | 91 |
| 40% | 0.68 | 7.05 | 0.10 | 314 | 0.83 | 5.07 | 0.16 | 86 |
| 50% | 0.66 | 7.00 | 0.09 | 312 | 0.93 | 5.11 | 0.18 | 79 |
| 60% | 0.59 | 7.02 | 0.08 | 328 | 1.24 | 4.74 | 0.26 | 72 |
| 70% | 0.58 | 6.95 | 0.08 | 341 | 1.50 | 4.71 | 0.32 | 59 |
| 80% | 0.53 | 6.98 | 0.08 | 347 | 1.89 | 3.90 | 0.48 | 53 |
| 90% | 0.61 | 6.88 | 0.09 | 359 | 1.63 | 4.25 | 0.38 | 41 |

only that the underlying state is more likely to be calm during calm months. Hence even when the returns conditional on knowing the hidden state is normal, returns during months we classify as being calm or turbulent will be a mixture of two normals and exhibit excess kurtosis.

Table 11: SUMMARY STATS IN TURBULENT AND CALM MONTHS

Months with $\Pr(S_t = T|\mathcal{F}_{t-1})$ exceeding 50% are counted as *turbulent* and other months are counted as *calm.*

| FORECASTED STATE | MEAN | SD | SKEWNESS | KURTOSIS | MONTHS |
|---|---|---|---|---|---|
| PANEL A: IN SAMPLE | | | | | |
| CALM | 1.53 | 5.14 | -0.23 | 5.19 | 779/978 |
| TURBULENT | -0.47 | 14.54 | -1.73 | 8.57 | 199/978 |
| ALL | 1.12 | 8.03 | -2.47 | 21.04 | 978/978 |
| PANEL B: OUT OF SAMPLE | | | | | |
| CALM | 1.72 | 5.49 | -0.32 | 5.11 | 321/400 |
| TURBULENT | -0.01 | 13.05 | -1.10 | 5.18 | 79/400 |
| ALL | 1.22 | 7.64 | -1.53 | 10.94 | 400/400 |

Figure 4, plots the time series of the estimated predicted probability of the hidden state being turbulent in a given calendar month along with an indicator as to whether it is recession month according to NBER. As can be seen, there is not much of a relationship between the month being in an NBER recession and the associated probability of the state being turbulent.

We also examine the association between the probability of the state being turbulent in a given month and likelihood of a momentum crash during that month using the following Probit model:

$$\Pr\left(R_t^{\text{mom}} < \text{ THRESHOLD LOSS }\right) = \Phi\left(a + b\Pr\left(S_t = T|\mathcal{F}_{t-1}\right)\right) \tag{19}$$

where $\Phi$ is the CDF of standard normal distribution and THRESHOLD LOSS is a critical level that defines a momentum crash. This specification helps us evaluate whether $\Pr(S_t = T|\mathcal{F}_{t-1})$ is related to the left tail of momentum returns. We consider THRESHOLD LOSS $= -10\%, -12.5\%, -15\%, -17.5\%, -20\%$. Table 12 gives the estimated parameter values and the associated t-stats for the Probit model in equation (19) for the in-sample as

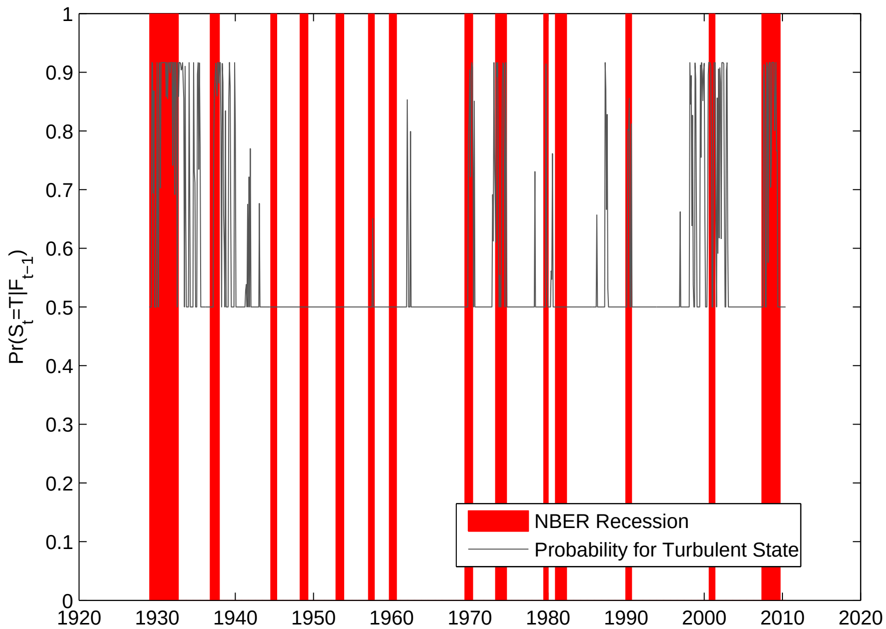


Figure 4: **NBER Recessions & Probability Of The Hidden State Being Turbulent**

well as the out-of-sample estimates of $\Pr(S_t = T|\mathcal{F}_{t-1})$. Except for the case when Threshold Loss $= -20\%$ for out-of-sample, $b$, the coefficient on $\Pr(S_t = T|\mathcal{F}_{t-1})$, is positive and statistically significant. The statistical insignificance for Threshold Loss $= -20\%$ out-of-sample, is probably due to there being too few months with large losses during the out-of-sample period. Interestingly, the size of $b$ monotonically increases as we lower Threshold Loss from $-10\%$ to $-20\%$, indicating that the probability of *more* extreme losses are *more* likely when $\Pr(S_t = T|\mathcal{F}_{t-1})$ is high.

Table 12: Probit Model of Momentum Crashes

| Threshold Loss | In Sample a | In Sample b | Out of Sample a | Out of Sample b |
|---|---|---|---|---|
| -10% | -2.11 | 1.40 | -1.89 | 1.16 |
| | (-19.08) | (7.41) | (-13.20) | (4.49) |
| -12.5% | -2.60 | 1.84 | -2.34 | 1.51 |
| | (-15.35) | (7.47) | (-11.44) | (4.81) |
| -15% | -2.67 | 1.84 | -2.42 | 1.40 |
| | (-14.62) | (7.03) | (-10.89) | (4.08) |
| -17.5% | -3.24 | 2.33 | -3.01 | 1.84 |
| | (-9.60) | (5.45) | (-7.34) | (3.53) |
| -20% | -5.04 | 4.13 | -7.14 | 6.14 |
| | (-3.20) | (2.29) | (-1.56) | (1.28) |

# 5 Alternative Specifications

In this section we examine a few alternative specifications for the stochastic process governing the temporal evolution of momentum returns and market excess returns. We relax the mean equation in the regime Switching model and let the beta of the momentum return to depend on past market conditions. We find that the more general specification is not necessarily better in terms of identifying months when large losses are more likely. We also evaluate a bivariate GARCH model of momentum and market excess returns.

For HMM, consider the following extended specification:

$$R_t^{\text{mom}} = \alpha(S_t) + \begin{pmatrix} \beta^0(S_t) & \\ + \ \beta^D(S_t) & \mathrm{I}_t^D \\ + \ \beta^U(S_t) & \mathrm{I}_t^U \\ + \ \beta^R(S_t) & \mathrm{I}_t^B \mathrm{I}_t^U \end{pmatrix} R_t^{\text{M}} + \sigma_{\text{mom}}(S_t)\epsilon_t^{\text{mom}}. \tag{20}$$

$$R_t^{\text{M}} = \mu\left(S_t\right) + \sigma_{\text{M}}\left(S_t\right)\epsilon_t^{M}, \tag{21}$$

Comparing to the equations (8) and (9), the above specification differs in that the beta of momentum portfolios can depend on past market conditions. We consider variations of the HMM specification by restricting the model parameters as given below:

| | |
|---|---|
| Main: | $\beta^R = 0, \beta^B = 0$ |
| Alt-1: | No Restriction |
| Alt-2: | $\beta^U = 0$ |
| Alt-3: | $\beta^U = 0, \beta^R = 0$ |
| Alt-4: | $\beta^U = 0, \beta^R = 0, \beta^B = 0$ |

Alt-1 is the most general specification. Alt-2 is similar to the specification in Daniel and Moskowitz (2011). Alt-3 captures the effects documented in Grundy and Martin (2001). Alt-4 corresponds to the market model. Note that Alt-2, Alt-3, Alt-4 and Main are nested within Alt-1. For $\mathrm{I}_t^B$ as in (6), we consider $L = 12, 24, 36$.

We also estimate the bivariate GARCH model given below:

$$R_t^{\mathrm{mom}} = \alpha + \left( \beta^0 + \beta^U \cdot \mathrm{I}_t^U \right) R_t^{\mathrm{M}} + \sigma_{\mathrm{mom,t}} \epsilon_t^{\mathrm{mom}}. \tag{22}$$

$$R_t^{\mathrm{M}} = \mu + \sigma_{\mathrm{M,t}} \epsilon_t^M, \tag{23}$$

where $\epsilon_t^{\mathrm{mom}}$ and $\epsilon_t^M$ are drawn from i.i.d standard normal distributions, and $\sigma_{\mathrm{mom,t}}$ and $\sigma_{\mathrm{M,t}}$ evolve according to the bivariate GARCH process given below:

$$\sigma^2_{\mathrm{mom,t}} = \sigma^2_{\mathrm{mom,0}} + p_1 \sigma^2_{\mathrm{mom,t-1}} + q_1 \left( \sigma_{\mathrm{mom,t-1}} \epsilon_{t-1}^{\mathrm{mom}} \right)^2 \tag{24}$$

$$\sigma^2_{\mathrm{M,t}} = \sigma^2_{\mathrm{M,0}} + p_2 \sigma^2_{\mathrm{M,t-1}} + q_2 \left( \sigma_{\mathrm{M,t-1}} \epsilon_{t-1}^{\mathrm{M}} \right)^2 \tag{25}$$

Table 13 gives ML estimates of the parameters for the various specifications given in (22), (23), (24) and (25).

Table 13: GARCH Model Parameter Estimates

| Mean-Return | | | GARCH of MOM | | | GARCH of MKT | | |
|---|---|---|---|---|---|---|---|---|
| | MLE | t-stat | | MLE | t-stat | | MLE | t-stat |
| $\alpha$ | 2.11 | 8.83 | $\sigma^2_{\mathrm{mom,0}}$ | 3.46 | 3.82 | $\sigma^2_{\mathrm{M,0}}$ | 0.69 | 2.99 |
| $\beta^0$ | 0.19 | 2.97 | $p_1$ | 0.63 | 11.05 | $p_2$ | 0.85 | 41.68 |
| $\beta^U$ | -0.55 | -5.56 | $q_1$ | 0.34 | 5.35 | $q_2$ | 0.13 | 5.98 |
| $\mu$ | 0.71 | 5.37 | | | | | | |

For the bivariate GARCH model, using the estimated parameter values, we compute $Std(R_t^{\text{mom}}|\mathcal{F}_{t-1})$, the conditional volatility of the momentum return for each month and use it as a measure of the tail-risk of the momentum returns.

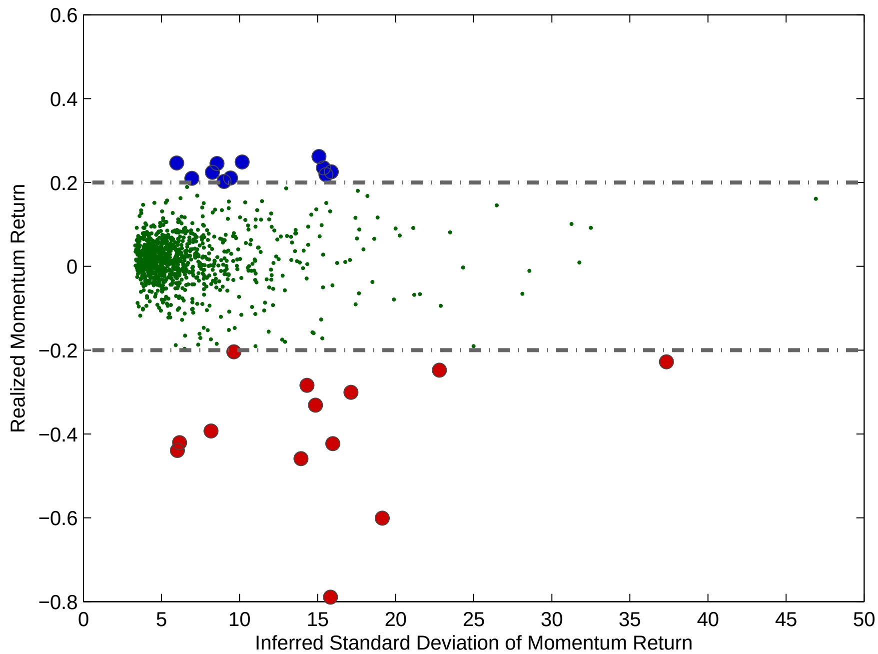


Figure 5: **Realized Momentum vs GARCH forecast of Volatility**

In Figure 5 we plot the realized momentum return in a month against its standard deviation according to the bivariate GARCH model. A comparison of Figure 5 with Figures 2 and 3 indicates that the association between momentum losses in a month and the riskiness of the month as measured by the standard deviation of momentum return during that month is not as strong for the bivariate GARCH model.

For comparing different models, we count the minimum number of months to be classified as *turbulent* such that all months with large losses (exceeding a threshold) occur during those *turbulent* months. Let $i_t$ denote the predicted level of tail-risk in a month $t$: $i_t \equiv \Pr(S_t = T|\mathcal{F}_{t-1})$ for HMM and $i_t \equiv Std(R_t^{\text{mom}}|\mathcal{F}_{t-1})$ for GARCH. For each threshold level of large loss $\underline{L}$, we pick an $\underline{i(L)}$ such that months exceeding the threshold level of loss $\underline{L}$ occur during months with $i_t > \underline{i(L)}$ that are classified as being turbulent. We then count the number of *turbulent* months corresponding to each threshold loss $\underline{L}$. We view the model that require classifying fewer number of months as being *turbulent* to capture months with large momentum losses as better.

Table 14 reports the number of months, the negative of a measure of a model to identify months when large losses are more likely. For extreme losses exceeding 20%, HMM specifications are much more effective than GARCH. Among HMM models, for $L = 12$, Alt-1 and Alt-2 require smaller number of months than the Main Model, but the differences are small. For out-of-sample months, the Main Model classifies 40 months as turbulent months, in contrast to 87 months classified as turbulent months with GARCH model. Also, among various HMM specifications, the Main HMM performs as well or better.

In Section 4, we observed that the relation between the probability of large gains and the probability of the hidden state being turbulent was weaker than the corresponding relation for large losses. We therefore compare the sample mean of momentum returns during months for turbulent and calm months. If an indicator for the momentum risk can detect large losses only, the sample mean will be significantly negative for turbulent months.

The properties of momentum returns during turbulent months – i.e., months when the probability for the hidden state being turbulent(HMM criterion) or conditional standard deviation of momentum returns(GARCH criterion) are sufficiently high to identify all months with losses exceeding a given threshold level – are reported in table 15. Interestingly, with a -20% threshold loss, with HMM, *less* months are classified as turbulent; the sample mean of momentum returns during turbulent months is *lower*; and the sample standard deviation of momentum returns during turbulent months is *higher*, when compared to the GARCH model.

For out-of-sample months, the differences becomes more significant. With the threshold loss of -20%, *less* than a half months are classified as turbulent with HMM when compared to GARCH. The sample mean of momentum returns during the 40 turbulent months is -3.51% per month for HMM. With the GARCH model, the average momentum return is -0.11% per month during the 87 months classified as being turbulent. Furthermore, as we lower the level of threshold loss from -10% to -20%, the standard deviation of momentum returns during turbulent months increases much more for HMM when compared to GARCH.

Table 14: **Comparison of Models: Number of Turbulent Months to Capture Large Losses**

| Threshold Loss | Model | In Sample | | | | Out of Sample | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | L=12 | L=24 | L=36 | | L=12 | L=24 | L=36 |
| -10% | Main | 941 | | | | 381 | | | |
| | Alt-1 | | 965 | 959 | 964 | | 389 | 384 | 388 |
| | Alt-2 | | 970 | 960 | 962 | | 382 | 388 | 385 |
| | Alt-3 | 970 | | | | 383 | | | |
| | Alt-4 | 951 | | | | 382 | | | |
| | GARCH | 942 | | | | 396 | | | |
| -12.5% | Main | 570 | | | | 205 | | | |
| | Alt-1 | | 497 | 495 | 499 | | 201 | 201 | 204 |
| | Alt-2 | | 506 | 464 | 484 | | 223 | 199 | 210 |
| | Alt-3 | 557 | | | | 250 | | | |
| | Alt-4 | 601 | | | | 246 | | | |
| | GARCH | 480 | | | | 185 | | | |
| -15% | Main | 368 | | | | 147 | | | |
| | Alt-1 | | 378 | 378 | 390 | | 147 | 152 | 149 |
| | Alt-2 | | 399 | 379 | 388 | | 164 | 150 | 162 |
| | Alt-3 | 422 | | | | 173 | | | |
| | Alt-4 | 420 | | | | 179 | | | |
| | GARCH | 480 | | | | 157 | | | |
| -17.5% | Main | 292 | | | | 123 | | | |
| | Alt-1 | | 274 | 276 | 275 | | 125 | 119 | 122 |
| | Alt-2 | | 286 | 283 | 281 | | 140 | 122 | 127 |
| | Alt-3 | 300 | | | | 145 | | | |
| | Alt-4 | 323 | | | | 160 | | | |
| | GARCH | 480 | | | | 133 | | | |
| -20% | Main | 148 | | | | 40 | | | |
| | Alt-1 | | 147 | 153 | 156 | | 51 | 52 | 54 |
| | Alt-2 | | 145 | 184 | 156 | | 50 | 49 | 52 |
| | Alt-3 | 176 | | | | 48 | | | |
| | Alt-4 | 154 | | | | 41 | | | |
| | GARCH | 236 | | | | 87 | | | |

Table 15: HMM vs GARCH: Number of Turbulent Months to Capture Large Losses (In Sample)

| | HMM | | | GARCH | | |
|---|---|---|---|---|---|---|
| Threshold Loss | # of Months | Mean | Std | # of Months | Mean | Std |
| -10% | 941 | 1.11 | 8.16 | 942 | 1.08 | 8.16 |
| -12.50% | 570 | 0.97 | 9.89 | 480 | 0.60 | 10.56 |
| -15% | 368 | 0.19 | 11.64 | 480 | 0.60 | 10.56 |
| -17.50% | 292 | -0.05 | 12.65 | 480 | 0.60 | 10.56 |
| -20% | 148 | -0.92 | 15.26 | 236 | -0.10 | 13.37 |

Table 16: HMM vs GARCH: Number of Turbulent Months to Capture Large Losses (Out of Sample)

| | HMM | | | GARCH | | |
|---|---|---|---|---|---|---|
| Threshold Loss | # of Months | Mean | Std | # of Months | Mean | Std |
| -10% | 381 | 1.25 | 7.78 | 396 | 1.23 | 7.65 |
| -12.50% | 205 | 0.86 | 9.71 | 185 | 0.43 | 10.12 |
| -15% | 147 | 0.45 | 10.83 | 157 | 0.58 | 10.57 |
| -17.50% | 123 | 0.42 | 11.34 | 133 | 0.66 | 11.02 |
| -20% | 40 | -3.51 | 15.22 | 87 | 0.11 | 11.98 |

# 6 Conclusion

Relative strength strategies, also known as momentum strategies are widely used by active quantitative portfolio managers and individual investors. These strategies generate large positive returns on average with little systematic risk as measured using standard asset pricing models and remain an anomaly.

In this paper we studied the returns on one such momentum strategy. During the 978 months covering July 1929 - December 2010 the returns on that widely studied strategy using U.S. stocks generated an average monthly return in excess of 1.12%/month and an alpha of 1.70%/month with respect to the three Fama and French (1993) factor model. Momentum strategy returns, when combined with the Fama and French three economy wide pervasive factor returns, gives rise to a portfolio with a Sharpe Ratio of almost 0.28 per month.

However momentum strategies also incur infrequent but rather large losses. There were 13 months with losses exceeding 20%/month in the sample of 978 months. The probability of such an event occurring if momentum strategy returns were independently and normally distributed would be 0.04%. We show that such periodic but rare large loss episodes can be captured by a two state hidden Markov model, where one state is turbulent and the other is calm. We find that it is possible to predict which of the two hidden state the economy is in with some degree of confidence. All the 13 months with losses exceeding 20%/month occur during turbulent months, *i.e.*, months when the predicted probability of the hidden state being turbulent exceeds 0.5. The probability of 13 months with losses exceeding 20% increases to 60% and momentum losses are less of a Black Swan. Momentum returns averaged -0.47%/month during turbulent months, with a Sharpe Ratio of -0.03. When such turbulent states are avoided, the monthly Sharpe Ratio of momentum strategy returns increases to 0.30 and price momentum poses still more of a challenge to standard asset pricing models.

# A Non-normality of distributions of residuals

We estimated the parameters of the hidden Markov model by maximizing the likelihood function assuming that the residuals in equations (8) and (9) are drawn from an i.i.d bivariate normal distribution. In what follows we show that the consistency of the estimator does not depend on i.i.d bivariate normality, so long as the expectation of the score function is zero at the true parameter values, i.e.,

$$\mathbb{E}\left[h\left(y_t;\theta\right)\right]=0 \tag{26}$$

where $h\left(y_t;\theta\right)=\frac{\partial \log \Pr(y_t|\mathcal{F}_{t-1})}{\partial\theta}$.

For notational convenience, we will in general not differentiate $S_t$, the random hidden state, from its realization, $s_t$.

We need the following assumptions for (26) to hold; it is not necessary that the residuals in equations (8) and (9) are drawn from an i.i.d bivariate normal distribution.

$$\begin{aligned}
\mathbb{E}\left[\epsilon_t^{\mathrm{mom}}\left(S_t\right)|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 0 & (27)\\
\mathbb{E}\left[R_t^{\mathrm{M}}\epsilon_t^{\mathrm{mom}}\left(S_t\right)|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 0 & (28)\\
\mathbb{E}\left[I_t^{U}R_t^{\mathrm{M}}\epsilon_t^{\mathrm{mom}}\left(S_t\right)|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 0 & (29)\\
\mathbb{E}\left[\left(\epsilon_t^{\mathrm{mom}}\left(S_t\right)\right)^2|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 1 & (30)\\
\mathbb{E}\left[\epsilon_t^{\mathrm{M}}|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 0 & (31)\\
\mathbb{E}\left[\left(\epsilon_t^{\mathrm{M}}\left(S_t\right)\right)^2|S_t,S_{t-1},\mathcal{F}_{t-1}\right] &= 1 & (32)\\
\mathbb{E}\left[I\left(S_t=S_{t-1}\right)|S_{t-1}=C,\mathcal{F}_{t-1}\right] &= p & (33)\\
\mathbb{E}\left[I\left(S_t=S_{t-1}\right)|S_{t-1}=T,\mathcal{F}_{t-1}\right] &= q & (34)
\end{aligned}$$

where $\epsilon_t^{\mathrm{mom}}$ and $\epsilon_t^{\mathrm{M}}$ are defined as in (13) and (14). Equations (27) - (32) will be satisfied when (8) and (9) are conditional regressions, i.e., the residuals are orthogonal to the right side variables. Equations (33) and (34) hold by assumption that the transition probabilities depend only on the current state.

From (10), (11), and (15), we can write $\log\Pr\left(y_t|\mathcal{F}_{t-1}\right)$ as follows:

$$\log\Pr\left(y_t|\mathcal{F}_{t-1}\right)=\log\sum_{S_t}\sum_{S_{t-1}}\Pr\left(y_t|S_t\right)\Pr\left(S_t|S_{t-1}\right)\Pr\left(S_{t-1}|\mathcal{F}_{t-1}\right). \tag{35}$$

In addition, from the expression of (12), the conditional log likelihood of realizations of observable random variables, $y_t$, is written as:

$$\log\Pr\left(y_t|S_t\right)=-\log\sigma_{\text{mom}}\left(S_t\right)-\frac{\left(\epsilon_t^{\text{mom}}\left(S_t\right)\right)^2}{2}-\log\sigma_{\text{M}}\left(S_t\right)-\frac{\left(\epsilon_t^{\text{M}}\left(S_t\right)\right)^2}{2}+C, \tag{36}$$

where $C$ is a constant.

First, let us consider the elements of $h\left(y_t;\theta\right)$ that correspond to first derivative the conditional log likelihood function with respect to one of the following parameters:

$$\alpha\left(S_t\right),\beta^0\left(S_t\right),\beta^U\left(S_t\right),\sigma_{\text{mom}}\left(S_t\right),\mu\left(S_t\right),\sigma_{\text{M}}\left(S_t\right).$$

Since

$$\mathbb{E}\left[h\left(y_t;\theta\right)\right]=\mathbb{E}\left[\mathbb{E}\left[h\left(y_t;\theta\right)|S_t,S_{t-1},\mathcal{F}_{t-1}\right]\right],$$

it is sufficient to show that

$$\mathbb{E}\left[h\left(y_t;\theta\right)|S_t,S_{t-1},\mathcal{F}_{t-1}\right]=0$$

for equation (26) to hold. Note that $\Pr\left(S_t|S_{t-1}\right)\Pr\left(S_{t-1}|\mathcal{F}_{t-1}\right)$ in (35) becomes 1 when we condition on the information set, $\{S_t,S_{t-1},\mathcal{F}_{t-1}\}$.

By differentiating the log likelihood function with respect to $\alpha\left(S_t\right)$ we get:

$$\frac{\partial\log\Pr\left(y_t|S_t\right)}{\partial\alpha\left(S_t\right)}=\epsilon_t^{\text{mom}}\left(S_t\right)\frac{1}{\sigma_{\text{mom}}\left(S_t\right)}.$$

Using (27), we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_t, S_{t-1}, \mathcal{F}_{t-1}\right)}{\partial \alpha\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\epsilon_t^{\text{mom}}\left(S_t\right)\frac{1}{\sigma_{\text{mom}}\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = 0 \tag{37}$$

Differentiating the log likelihood function with respect to $\beta^0\left(S_t\right)$ gives:

$$\frac{\partial \log \Pr\left(y_t|S_t\right)}{\partial \beta^0\left(S_t\right)} = \epsilon_t^{\text{mom}}\left(S_t\right)\frac{R_t^{\text{M}}}{\sigma_{\text{mom}}\left(S_t\right)}.$$

Using (28) we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_t, S_{t-1}, \mathcal{F}_{t-1}\right)}{\partial \beta^0\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\epsilon_t^{\text{mom}}\left(S_t\right)\frac{R_t^{\text{M}}}{\sigma_{\text{mom}}\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = 0 \tag{38}$$

Differentiating the log likelihood function with respect to $\beta^U\left(S_t\right)$ gives:

$$\frac{\partial \log \Pr\left(y_t|S_t\right)}{\partial \beta^U\left(S_t\right)} = \epsilon_t^{\text{mom}}\left(S_t\right)\frac{I_t^U R_t^{\text{M}}}{\sigma_{\text{mom}}\left(S_t\right)}.$$

Using (29) we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_t, \mathcal{F}_{t-1}\right)}{\partial \beta^0\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\epsilon_t^{\text{mom}}\left(S_t\right)\frac{I_t^U R_t^{\text{M}}}{\sigma_{\text{mom}}\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = 0 \tag{39}$$

Differentiating the log likelihood function with respect to $\sigma_{\text{mom}}\left(S_t\right)$ gives:

$$\frac{\partial \log \Pr\left(y_t|S_t\right)}{\partial \sigma_{\text{mom}}\left(S_t\right)} = -\frac{1}{\sigma_{\text{mom}}\left(S_t\right)} + \frac{1}{\sigma_{\text{mom}}\left(S_t\right)}\left(\epsilon_t^{\text{mom}}\left(S_t\right)\right)^2.$$

Using (30) we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_t\right)}{\partial \sigma_{\text{mom}}\left(S_t\right)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[-\frac{1}{\sigma_{\text{mom}}\left(S_t\right)} + \frac{1}{\sigma_{\text{mom}}\left(S_t\right)}\left(\epsilon_t^{\text{mom}}\left(S_t\right)\right)^2|S_t, S_{t-1}\mathcal{F}_{t-1}\right] = 0 \tag{40}$$

Differentiating the log likelihood function with respect to $\mu(S_t)$ gives:

$$\frac{\partial \log \Pr(y_t|S_t)}{\partial \mu(S_t)} = \epsilon_t^{\mathrm{M}}(S_t)$$

Using (31) we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr(y_t|S_t)}{\partial \mu(S_t)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[\epsilon_t^{\mathrm{M}}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = 0 \tag{41}$$

Differentiating the log likelihood function with respect to $\sigma_{\mathrm{M}}(S_t)$ gives:

$$\frac{\partial \log \Pr(y_t|S_t)}{\partial \sigma_{\mathrm{M}}(S_t)} = -\frac{1}{\sigma_{\mathrm{M}}(S_t)} + \frac{1}{\sigma_{\mathrm{M}}(S_t)}\left(\epsilon_t^{\mathrm{M}}(S_t)\right)^2.$$

Using (32) we get:

$$\mathbb{E}\left[\frac{\partial \log \Pr(y_t|S_t)}{\partial \sigma_{\mathrm{M}}(S_t)}|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = \mathbb{E}\left[-\frac{1}{\sigma_{\mathrm{M}}(S_t)} + \frac{1}{\sigma_{\mathrm{M}}(S_t)}\left(\epsilon_t^{\mathrm{M}}\right)^2|S_t, S_{t-1}, \mathcal{F}_{t-1}\right] = 0 \tag{42}$$

Thus, the condition of (26) is satisfied for $\alpha(S_t), \beta^0(S_t), \beta^U(S_t), \sigma_{\mathrm{mom}}(S_t), \mu(S_t), \sigma_{\mathrm{M}}(S_t)$ with the results from (37) to (42), implied by the assumptions from (27) to (32).

Next, we consider the elements of the score that correspond to differentiating the log likelihood function with respect to the parameters $p$ and $q$. Since

$$\mathbb{E}[h(y_t;\theta)] = \mathbb{E}[\mathbb{E}[h(y_t;\theta)|S_{t-1}, \mathcal{F}_{t-1}]],$$

it is sufficient to show that

$$\mathbb{E}[h(y_t;\theta)|S_{t-1}, \mathcal{F}_{t-1}] = 0$$

for the condition of (26) to hold. Conditional on the information set of $\{S_{t-1}, \mathcal{F}_{t-1}\}$, we can treat $\Pr(S_{t-1}|\mathcal{F}_{t-1})$ in the expression of (35) as 1.

Consider a case that $S_{t-1} = C$. Then,

$$
\begin{aligned}
&\log \Pr\left(y_t | S_{t-1} = C, \mathcal{F}_{t-1}\right) \\
=& \log \sum_{S_t} \Pr\left(y_t | S_t\right) \Pr\left(S_t | S_{t-1} = C\right) \\
=& \log \left(\Pr\left(y_t | S_t = C\right) I(S_t = C) p + \Pr\left(y_t | S_t = T\right) I(S_t = T)\left(1 - p\right)\right).
\end{aligned}
\tag{43}
$$

Thus,

$$
\begin{aligned}
&\frac{\partial \log \Pr\left(y_t | S_{t-1} = C, \mathcal{F}_{t-1}\right)}{\partial p} \\
=& \frac{\Pr\left(y_t | S_t = C\right) I(S_t = C) - \Pr\left(y_t | S_t = T\right) I(S_t = T)}{\Pr\left(y_t | S_t = C\right) I(S_t = C) p + \Pr\left(y_t | S_t = T\right)\left(1 - p\right) I(S_t = T)} \\
=& \frac{1}{p} I\left(S_t = C\right) - \frac{1}{1-p} I\left(S_t = T\right)
\end{aligned}
\tag{44}
$$

Using (33) we get:

$$
\begin{aligned}
&\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t | S_{t-1} = C, \mathcal{F}_{t-1}\right)}{\partial p} | S_{t-1} = C, \mathcal{F}_{t-1}\right] \\
=& \mathbb{E}\left[\frac{1}{p} I\left(S_t = C\right) - \frac{1}{1-p} I\left(S_t = T\right) | S_{t-1} = C, \mathcal{F}_{t-1}\right] \\
=& \frac{p}{p} - \frac{1-p}{1-p} = 0.
\end{aligned}
\tag{45}
$$

Furthermore, since $q$ is not related to the conditional likelihood in a case that $S_{t-1} = C$, it is trivial to show that

$$
\frac{\partial \log \Pr\left(y_t | S_{t-1} = C, \mathcal{F}_{t-1}\right)}{\partial q} = 0, \tag{46}
$$

implying that

$$
\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t | S_{t-1} = C, \mathcal{F}_{t-1}\right)}{\partial q} | S_{t-1} = C, \mathcal{F}_{t-1}\right] = 0. \tag{47}
$$

Similarly, we can show the followings hold:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_{t-1}=T,\mathcal{F}_{t-1}\right)}{\partial p}|S_{t-1}=T,\mathcal{F}_{t-1}\right]=0 \tag{48}$$

and

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|S_{t-1}=T,\mathcal{F}_{t-1}\right)}{\partial q}|S_{t-1}=T,\mathcal{F}_{t-1}\right]=0 \tag{49}$$

From (45), (47), (48), and (49), it follows that:

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|\mathcal{F}_{t-1}\right)}{\partial p}\right]=0$$

and

$$\mathbb{E}\left[\frac{\partial \log \Pr\left(y_t|\mathcal{F}_{t-1}\right)}{\partial q}\right]=0$$

That is, the condition of (26) holds for $p, q$.

We have shown that the expectation of score function is zero without relying on normality. Now we define GMM estimator using the moment condition, (26).

$$\theta_{GMM}=\arg\min_{\theta\in\Theta} g_T' g_T \tag{50}$$

where $\Theta$ is an compact set such that $\theta^0\in\Theta$ and $g_T$ is defined as follows:

$$g_T=\frac{1}{T}\sum_{t=1}^{T} h\left(y_t;\theta\right) \tag{51}$$

Note that the GMM estimator defined in (50) is identical to the ML estimator defined in (18).

For the consistency of parameter estimates, we need an additional condition that only the true parameters satisfy the moment condition of (26). Due to the continuity and differentiability of the score function, the consistency and the asymptotic normality of the estimator defined in (50) directly follow from the general results for GMM estimators, as summarized in the following propositions:

**Proposition 1** *Assume that only $\theta^0$ satisfies equations (26). As $T\rightarrow\infty$, the estimator* *defined in (50) converges in probability to $\theta^0$.*

**Proposition 2** *Assume that only $\theta^0$ satisfies equation (26). As $T \to \infty$, $\sqrt{T}\left(\theta_{GMM} - \theta^0\right)$ converges in distribution to $N(0, V)$ where $V = \left(DS^{-1}D'\right)^{-1}$ and $D$ is the probability limit of $\frac{\partial g_T}{\partial \theta'}$ and $S$ is the asymptotic variance of $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} h\left(y_t; \theta\right)$*

In the computation of $\hat{V} = \left(\hat{D}\hat{S}^{-1}\hat{D}'\right)^{-1}$ in section 4, we set $\hat{D}$ as $\frac{1}{T}$ times the hessian of log-likelihood function and $\hat{S}$ as $\frac{1}{T} \sum_{t=1}^{T} h\left(y_t; \hat{\theta}\right) h\left(y_t; \hat{\theta}\right)'$ where $h\left(y_t; \hat{\theta}\right)$ is the numerical derivative of $\log \Pr\left(y_t | \mathcal{F}_{t-1}\right)$ at $\hat{\theta}$.

# References

Ambasta, M, and Arik Ben Dor, 2010, Barclays capital alternative replicators: A synthetic approach to replicating hedge fund performance, *Barclays Capital Report*.

Andersen, Torben G., Oleg Bondarenko, and Maria T. Gonzalez-Perez, 2011, Coherent model-free implied volatility: A corridor fix for high-frequency vix, CREATES Research Papers 2011-49 School of Economics and Management, University of Aarhus.

Ang, Andrew, and Allan Timmerman, 2011, Regime changes and financial markets, Working Paper.

Asness, Clifford S., Tobias Moskowitz, and Lasse Pedersen, 2008, Value and momentum everywhere, University of Chicago working paper.

Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, 2011, Conditional risk and performance evaluation: Volatility timing, overconditioning, and new estimates of momentum alphas, *Journal of Financial Economics* 102, 363–389.

Bollerslev, Tim, and Viktor Todorov, 2011, Tails, fears and risk premia, *Journal of Finance* 66, 2165–2211.

Breen, William, Lawrence R. Glosten, and Ravi Jagannathan, 1989, Economic signficance of predictable variation in stock index returns, *Journal of Finance* 44, 1177 – 1189.

Calvet, Laurent E., and Adlai J. Fisher, 2008, *Multifractal Volatility: Theory, Forecasting, and Pricing* (Academic Press: London, UK).

Chabot, Benjamin, Eric Ghysels, and Ravi Jagannathan, 2009, Momentum cycles and limits to arbitrage - evidence from victorian england and post-depression us stock markets, Working Paper.

Chan, K.C., 1988, On the contrarian investment strategy, *Journal of Business* 61, 147–163.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, Market states and momentum, *Journal of Finance* 59, 1345–1365.

Daniel, Kent D., and Tobias J. Moskowitz, 2011, Momentum crashes, Columbia Business School working paper.

DeBondt, Werner F. M., and Richard H. Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557–581.

Elavia, Tony, and Migene Kim, 2011, Managing the volatility of alpha models, *Journal of Investment Management* 9, 1–16.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F. Fama, and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653–1678.

Glosten, Lawerence R., and Ravi Jagannathan, 1994, A contingent claim approach to performance evaluation, *Journal of Empirical Finance* 1, 133–160.

Griffin, John M., Xiuqing Ji, and J. Spencer Martin, 2003, Momentum investing and business cycle risk: Evidence from pole to pole, *Journal of Finance* 58, 2515–2547.

Grinblatt, Mark, and Sheridan Titman, 1989, Mutual fund performance: an analysis of quarterly portfolio holdings, *Journal of Business* 62, 393–416.

——— , 1993, Performance measurement without benchmarks: An examination of mutual fund returns, *Journal of Business* 66, 47–68.

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

Hamilton, James D, 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* 57, 357–84.

Hansen, Lars P., and Ravi Jagannathan, 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* 99, 225–262.

Henriksson, Roy D., and Robert C. Merton, 1981, On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, *Journal of Business* 54, 513–533.

Jagannathan, Ravi, and Robert A Korajczyk, 1986, Assessing the market timing performance of managed portfolios, *Journal of Business* 59, 217–35.

Jagannathan, Ravi, and Zhenyu Wang, 1996, The conditional CAPM and the cross-section of expected returns, *Journal of Finance* 51, 3–53.

Jegadeesh, Narasimhan, 1990, Evidence of predictable behavior of security returns, *Journal of Finance* 45, 881–898.

——— , and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

——— , 2005, *Momentum, in Richard Thaler, ed.: Advances in Behavioral Finance II* (Princeton Press: Princeton, NJ).

Jensen, Michael C., 1967, Random walks: Reality or myth: Comment, *Financial Analysts Journal* 23, 77–85.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs?, *Journal of Finance* 59, 1039–1082.

Kothari, S.P., and Jay Shanken, 1992, Stock return variation and expected dividends, *Journal of Financial Economics* 31, 177–210.

Lehmann, Bruce N., 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* 105, 1–28.

Levy, Robert A., 1967, Relative strength as a criterion for investment selection, *The Journal of Finance* 22, 595–610.

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

Swaminathan, Bhaskaran, 2010, Qunatitative money management: A practical application of behavioral finance, *Working Paper*.

Todorov, Viktor, and George Tauchen, 2011, Volatility jumps, *Journal of Business & Economic Statistics* 29, 356–371.