A Hidden Markov Model of Leverage Dynamics, Tail Risk, and Value-Momentum Correlation

Kent Daniel, Ravi Jagannathan and Soohun Kim[†]

Abstract

Momentum strategies exhibit rare but dramatic losses (crashes), which we show are a result of the leverage dynamics of stocks in the momentum portfolio. When the economy is in a hidden *turbulent* state, associated with a depressed and volatile stock market, the short-side of the momentum portfolio becomes highly levered, and behaves like a call option on the market index portfolio, making momentum crashes more likely. We develop a hidden Markov model of the unobserved *turbulent* state that affects the joint returns of the momentum strategy and the market index portfolios. We find that the use of a combination of normal and Student's t-distributions for the hidden residuals in the model to construct the likelihood of the realized momentum and market index returns dramatically improves the model's ability to predict crashes. The same variable that forecasts momentum crashes also forecasts the correlation between momentum and value, two of the benchmark investment styles used in performance appraisal of quant portfolio managers. The correlation is negative only when the probability of the economy being in a *turbulent* state is high. The conditional correlation is zero otherwise, which is two thirds of the time. Half of the negative relationship between value and momentum is due to the leverage dynamics of stocks in the momentum portfolio. The other half is due to a hidden risk factor, likely related to funding liquidity identified in Asness et al. (2013), which emerges only when the economy is more likely to be in the turbulent state.

[†]Kent Daniel, Finance and Economics Division, Graduate School of Business, Columbia University and NBER; E-mail: kd2371@columbia.edu,
Ravi Jagannathan, Finance Department, Kellogg School of Management, Northwestern University and NBER, ISB, SAIF; E-mail: rjaganna@kellogg.northwestern.edu,
Soohun Kim, Finance Area, Georgia Institute of Technology; E-mail: soohun.kim@scheller.gatech.edu.

1 Motivation

Price momentum can be described as the tendency of securities with relatively high (low) past returns to subsequently outperform (underperform) the broader market. Long-short momentum strategies exploit this pattern by taking a long position in past winners and an offsetting short position in past losers. Momentum strategies have been, and continue to be popular among traders. A majority of quantitative fund managers employ momentum as a component of their overall strategy, and even fundamental managers appear to incorporate momentum in formulating their trading decisions.[1]

Notwithstanding their inherent simplicity, momentum strategies have been profitable across many asset classes and in multiple geographic regions.[2] Over our sample period of 1044 months from 1927:01 to 2013:12, our baseline momentum strategy produced monthly returns with a mean of 1.18% and a standard deviation of 7.94%, generating an annualized Sharpe ratio of 0.52.[3] Over this same period, the market excess return (Mkt-Rf) had an annualized Sharpe Ratio of 0.41. Momentum's CAPM alpha is 1.52%/month ($t=7.10$).[4]

While the momentum strategy's average risk adjusted return has been high, the strategy has experienced infrequent but large losses. As Chabot et al. (2014) document, this phe-

---

[1]Swaminathan (2010) shows that most quantitative managers make use of momentum. He 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. In addition 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." Badrinath and Wahal (2002) show that institutions behave as momentum traders when entering new positions, but as contrarians otherwise.

[2]Asness et al. (2013) provide extensive cross-sectional evidence for momentum effects. Chabot et al. (2014) find the momentum effect in the Victorian era UK equity market.

[3]Our baseline 12-2 momentum strategy, described in more detail later, ranks firms based on their cumulative returns from months $t-12$ through $t-2$, and takes a long position in the value-weighted portfolio of the stocks in the top decile, and a short position in the value-weighted portfolio of the bottom decile stocks.

[4]Over the same period, the SMB and HML factors by Fama and French (1993) had annualized Sharpe Ratios of 0.26 and 0.39, respectively, and the Fama and French three-factor alpha for momentum is 1.76%/month ($t=8.20$). From 1967:01 to 2013:12, the I/A and ROE factors by Hou et al. (2015) achieved annualized Sharpe Ratios of 0.81 and 0.77, respectively, the Hou et al. (2015) four-factor alpha for momentum is 0.39%/month ($t=1.07$). Lastly, the annualized Sharpe Ratios of CMW and RMA factors in Fama and French (2015) are 0.41 and 0.57, respectively, and the associated five-factor alpha for momentum is 1.34%/month ($t=4.03$). The t-statistics are computed using the heteroskedasticity-consistent covariance estimator by White (1980).

nomenon is robust.[5] The historical distribution of momentum strategy returns is highly left skewed. Consistent with the large estimated negative skewness, over our sample, there are eight months in which the momentum strategy lost more than 30%, and none in which it has earned more than 30% (the highest monthly return is 26.18%). Moreover, the strategy's largest losses have been extreme. The worst monthly return was -79.57%, and six monthly losses exceed 40%. Normality can easily be rejected. Also, as Daniel and Moskowitz (2016) document, these large losses cluster, and tend to occur when the market rebounds sharply following a prolonged depressed condition.

The focus of this paper is modeling time variation in the tail risk of momentum strategies. We argue that the way momentum strategy portfolios are constructed necessarily embeds a written call option on the market portfolio, with time varying moneyness. The intuition comes from Merton (1974): following large negative market returns, the effective leverage of the firms on the short side of the momentum strategy (the past-loser firms) becomes extreme. As the firm value falls, the common shares move from being deep in-the-money call options on the firm's underlying assets, to at- or out-of-the-money options, and thus start to exhibit the convex payoff structure associated with call options: the equity value changes little in response to even large down moves in the underlying firm value, but moves up dramatically in response to large up moves. Thus, when the values of the firms in the loser portfolio increase—proxied by positive returns on the market portfolio—the convexity in the option payoff results in outsized gains in the past loser portfolio. Since the momentum portfolio is short these loser firms, this results in dramatic losses for the overall long-short momentum portfolio.

Interestingly, this same apparent optionality is observed not only in cross-sectional equity momentum strategies, but also in commodity and currency momentum strategies (see Daniel

---

[5]Using self-collected UK historical data from 1867 to 1907 and CRSP data from 1926 to 2012, the authors find that, while momentum has earned abnormally high risk-adjusted returns, the strategy also exposed investors to large losses (crashes) during both periods.

and Moskowitz (2016)). Related arguments suggest that effective leverage is likely to be the driver of this same optionality: in the case of commodity momentum, this option-like feature likely arises from the lower bound on variable costs associated with production, the option to shut down, and the lead times involved in adjusting production. In the case of currencies, central bankers tend to lean against the wind and when that effort fails, currency prices tend to crash. Further, those who engage in currency carry trades borrow in the low interest rate currency and lend in the high interest rate currency. When interest rate differentials change, sudden unwinding of large currency trade positions due to margin calls can lead to large FX momentum strategy losses. For example, during 1997-98, the US dollar interest rates were higher than the Yen interest rates, and the US dollar was steadily appreciating against the Yen before crashing in October 1998. One explanation for the sharp rise of the Yen against the US dollar is the drop in US interest rates and the sudden unwinding of Dollar-Yen carry trade positions by hedge funds with weaker capital positions from exposure to 1998 Russian crisis.[6]

For US common stocks, we show that the dynamics of reported financial leverage are consistent with this hypothesis: going into the five worst momentum crash months, financial leverage of the loser portfolio averaged 47.2, more than an order of magnitude higher than unconditional average of 3.97.[7] However, a firm's financial leverage is not a good proxy for that firm's *effective* leverage: firms have many fixed costs distinct from the repayment of their debt, including the wages of crucial employees, the fixed costs associated with maintenance of property, plant and equipment, etc. If these fixed costs are large, even a firm with zero debt may see its equity start to behave like an out-of-the-money option following large losses. A recent episode that is consistent with the view that non-financial leverage can increase option like feature was the collapse of many "dot-com" firms in the 2000-2002 period, where large drops in the values of these firms did not lead to large increases in financial leverage,

---

[6]See 69th Annual Report of the Bank for International Settlements, page 107.

[7]These are the averages over the 1964-2013 period over which we have data on the book value of debt.

yet clearly affected the operating leverage of these firms.

Because it is difficult to directly measure the effective leverage—operating plus financial—of the firms that make up the short-side of the momentum portfolio, we instead estimate the leverage dynamics of the momentum portfolio using hidden Markov model that incorporates this optionality. In the model, we assume that the economy can be viewed as being in one of just two unobserved states, *calm* and *turbulent*. We develop a two-state hidden Markov model (HMM) where the momentum return generating process is different across the two states, and estimate the probability that the economy is in the unobserved turbulent state using maximum likelihood. Our HMM specification can be viewed as a parsimonious dynamic extension of the return generating process in Henriksson and Merton (1981) and Lettau et al. (2014). One striking finding is that, while the momentum returns themselves are highly left-skewed and leptokurtic, the residuals of the momentum return generating process coming out of our estimated HMM specification are Normally distributed.[8] A key component of the HMM specification is the embedded option on the market; by looking for periods in time where the optionality is stronger, we can better estimate whether a momentum "tail event" is more likely. Consistent with this, we find that the HMM-based estimate of the turbulent state probability forecasts large momentum strategy losses far better than alternative explanatory variables such as past market and past momentum returns and their realized volatilities or volatility forecasts from a GARCH model that can be viewed as realized volatility computed using all past observations with more weight to the immediate past.

Interestingly, we find that it is the incorporation of the optionality in the HMM that is key to the model's ability to forecast these tail events. A version of the HMM which incorporates all other model components (i.e., the volatilities and mean returns of the both the market and the momentum portfolios), but which does not include the optionality, is not

---

[8]In contrast, the market-returns residuals have a Student's t-distribution with 5 degrees of freedom. We account for this non-Normality in one our HMM specifications and find that accounting for this non-Normality substantially improves the performance of the model in forecasting tail-events.

as successful: the model without the optionality produces about 30% more false positives than the baseline HMM specification, suggesting that the historical convexity in the relation between the market and momentum portfolio allows better estimation of the turbulent state probability. Intuitively, increasing leverage in the past loser portfolio, identified by the HMM as an increase in the convexity of the momentum strategy returns, presages future momentum crashes.

The literature examining price momentum is vast. While the focus in this literature has been on documenting and explaining the strategy's high average returns[9] and unconditional risk exposures, a more recent literature has focused on characterizing the time variation in the moments. Barroso and Santa-Clara (2015) study the time-varying volatility in momentum strategy returns. Daniel and Moskowitz (2016) find that infrequent large losses to momentum strategy returns are pervasive phenomena — they are present in several international equity markets and commodity markets — and they tend to occur when markets recover sharply from prolonged depressed conditions. Grundy and Martin (2001) examine the time-varying nature of momentum strategy's exposure to standard systematic risk factors. In contrast to most of this literature, our focus here is on the strategy's tail risk. In particular, we show how this tail risk arises, model it with our HMM, estimate this model and show that it captures these important tail risks better than other forecasting techniques suggested by the literature.

Our findings also contribute to the literature characterizing hidden risks in dynamic portfolio strategies and the literature on systemic risk. For example, Mitchell and Pulvino (2001) find that merger arbitrage strategy returns have little volatility and are market neutral during most times. However the strategies effectively embed a written put option on the market, and consequently tend to incur large losses when the market depreciates sharply. When a number of investors follow dynamic strategies that have embedded options on the

---

[9]See Daniel et al. (1998), Barberis et al. (1998), Hong and Stein (1999), Liu and Zhang (2008) and Han et al. (2009) for examples.

market of the same type, crashes can be exacerbated with the potential to trigger systemic responses.

While our focus is in modeling systematic, stochastic, variations in the tail risk of momentum returns—which we find is due to its embedded option on the market like features—our findings also have implications for estimating the abnormal returns to the momentum strategy. It is well recognized in the literature that payoffs on self financing zero cost portfolios that have positions in options can exhibit spurious positive value (alpha) when alpha is computed using the market model or linear beta models in general.[10] We therefore calculate an option-adjusted abnormal performance for the momentum strategy. As might be anticipated, we find that the alpha of the momentum strategy is generally strongly positive and statistically significant. However, when the *ex-ante* turbulent state probability is sufficiently high—and there are several historical episodes where it is—the estimated alpha is negative and statistically significant.

Interestingly, we find that the same state variable that forecasts momentum crashes also forecasts time-variation in the correlation between momentum and value, two of the benchmark investment styles frequently used in fund manager performance evaluation.[11] It is well known that the correlation between value and momentum returns is negative (see, e.g., Asness et al., 2013). We show that in fact the conditional correlation is negative only when the probability of the economy being in a *turbulent* state is high, which happens about a third of the time. The remainder of the time, the correlation is approximately zero. We further decompose this conditional covariance in the turbulent state, and show that half is attributable to the leverage dynamics of the individual firms in the momentum portfolio. The other half is due to a hidden risk factor which emerges only when the economy is more likely to be in the turbulent state. We argue that this hidden risk factor is likely related to

---

[10]See Jagannathan and Korajczyk (1986) for an example.

[11]Koijen et al. (2009) show that investors can achieve economically significant gains by utilizing the features of short-run momentum or long-run value in stock returns.

funding liquidity measures identified in Asness et al. (2013).

The rest of this paper is organized as follows. In Section 2, we examine the various drivers of momentum crashes, and show that these arise as a result of the strong written call option-like feature embedded in momentum strategy returns in certain market conditions. In Section 3, we describe a hidden Markov model for momentum's return generating process that captures this feature of tail risk in momentum strategy returns. In Section 4, we show the ability of our hidden Markov model to predict momentum crashes. In Section 5, we evaluate the conditional alpha of momentum strategy returns based on the estimated parameters of our hidden Markov model and option market prices. In Section 6 we show that, in addition to forecasting the probability of momentum strategy "crashes," the *ex-ante* turbulent state probability reliably forecasts the value-momentum return correlation and covariance. Interestingly, when the turbulent state probability is low, the conditional value-momentum return correlation is zero. Section 7 concludes.

2   Momentum Crashes

In this section, we describe the return on a particular momentum strategy that we examine in detail in this paper. We show that the distribution of momentum strategy returns is heavily skewed to the left and significantly leptokurtic. We also find that the return on the momentum strategy is non-linearly related to the excess return of the market index portfolio. The nature of non-linear relationship depends on market conditions. This examination motivates the two-state model that we develop in Section 3.

2.1   Characteristics of Momentum Strategy Returns

Price momentum strategies have been constructed using variety of metrics. For this study we examine a cross-sectional equity strategy in US common stocks. Our universe consists

of all US common stocks in CRSP with sharecodes of 10 and 11 which are traded on the NYSE, AMEX or NASDAQ. We divide this universe into decile portfolios at the beginning of each month t based on each stock's "(12,2)" return: the cumulative return over the 11 month period from months $t-12$ through $t-2$.[12] Our decile portfolio returns are the market-capitalization weighted returns of the stocks in that past return decile. A stock is classified as a "winner" if its (12-2) return would place it in the top 10% of all NYSE stocks, and as a "loser" if its (12-2) return is in the bottom 10%. Most of our analysis will concentrate on the zero-investment portfolio "MOM" which is long the past-winner decile, and short the past-loser decile.

Panel A of Table 1 provides various statistics describing the empirical distribution of the momentum strategy return (MOM) and the three Fama and French (1993) factors.[13] Without risk adjustment, the momentum strategy earns an average return of 1.18%/month and an impressive annualized Sharpe Ratio of 0.52. Panels B and C show that after risk adjustment, the average momentum strategy return increases: its CAPM alpha is 1.52%/month ($t=7.10$) and its Fama and French (1993) three factor model alpha is 1.76%/month ($t=8.20$).[14] This is not surprising given the negative unconditional exposure of MOM to the three factors.

The focus of our study is the large, asymmetric losses of the momentum strategy: Panel A of Table 1 shows that the MOM returns are highly left-skewed and leptokurtic. Figure 1.A illustrates this graphically: we plot the smoothed empirical density for MOM returns (the dashed red line) and a normal density with the same mean and standard deviation. Overlayed on the density function plot are red dots that represent the 25 MOM returns that exceed 20% in absolute value (13 in the left tail and 12 in the right tail). Figure 1.B overlays the empirical density of market excess returns which are scaled to match the volatility of MOM

---

[12]The one month gap between the return measurement period and the portfolio formation date is done both to be consistent with the momentum literature, and to minimize market microstructure effects and to avoid the short-horizon reversal effects documented in Jegadeesh (1990) and Lehmann (1990).

[13]The Mkt-Rf, SMB and HML return data come from Kenneth French's database.

[14]The t-statistics are computed using the heteroskedasticity-consistent covariance estimator by White (1980).

returns over this sample period. The 20 Mkt-Rf* returns that exceed 20% in absolute value (11 in the left tail and 9 in the right tail) are represented by blue dots.

Consistent with the results in Table 1, Figure 1 reveals that both the market and momentum strategy are leptokurtic. However, Panel B in particular shows the strong left skewness of momentum. Again, one of the objectives of this paper is to show that this skewness is completely a result of the time-varying non-linear relationship between market and momentum returns caused by the time-varying leverage of firms in the loser portfolio. As a way of motivating our model, we next examine the influence of prevailing state variables on market conditions on momentum strategy returns.

To begin, Table 2 lays out the MOM returns in the 13 months when the MOM loss exceeded 20%, and various measures of market conditions that prevailed during the months. The first set of columns show that the large momentum strategy losses are generally associated with large gains on the past-loser portfolio rather than losses in the past-winner portfolio. During the 13 large loss months, the loser portfolio earned an average excess return of 45.69% whereas the winner portfolio earned only 6.32%. Interestingly, these loser portfolio gains are associated with large contemporaneous gains in the market portfolio, which earns an average excess return of 16.14% in these months. However, the table also shows that market return is strongly negative and volatile in the period leading up to the momentum crashes: the market is down, on average, by more than 37% in the three years leading up to these crashes, and the market volatility is almost three times its normal level in the year leading up to the crash.[15] Given the past losses and high market volatility, it is not surprising that the past-loser portfolio has suffered severe losses: the threshold (breakpoint) for a stock to be in the loser portfolio averaged -63.77% in these 13 months, about 2.7 times the average breakpoint. Thus, at the start of the crash months, stocks in the past-loser portfolio are likely very highly levered. Table 2 also shows that the average financial leverage (book value

---

[15]Realized volatility is computed as the square root of the sum of squared daily returns and expressed as annualized percentage.

of debt/market value of equity), during the 5 large loss months after 1964 (when our leverage data starts) is 47.2, more than an order of magnitude higher than the average leverage of the loser portfolio, 3.97.

To summarize, large momentum strategy losses generally have occurred in volatile bear markets, when the past-losers have lost a substantial fraction of their market value, and consequently have high financial leverage, and probably high operating leverage as well. Thus, following Merton (1974), the equity of these firms behave like an out-of-the-money call option on the underlying firm value which, in aggregate, is correlated with the market. Consequently when the market recovers sharply, the loser portfolio experiences outsized gains, resulting in the extreme momentum strategy losses we observe.

2.2 Time Varying Option-like Features of Momentum Strategy

Motivated by the evidence in the preceding Section, we here examine the time-variation in the call-option-like feature of momentum strategy returns. This serves as motivation for the two-state HMM model that we develop in Section 3.

In particular, we consider the following augmented return generating process, similar to that considered by Henriksson and Merton (1981) and others:[16]

$$R_{p,t}^e = \alpha_p + \beta_p^0 R_{\text{MKT},t}^e + \beta_p^+ \max\left(R_{\text{MKT},t}^e, 0\right) + \varepsilon_{p,t}, \tag{1}$$

---

[16]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 they 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 portfolio returns in non-US markets. Daniel and Moskowitz (2016) show that the optionality is time varying, and is particularly pronounced in high volatility down markets, and is driven by the behavior of the short-side (loser) as opposed to the long (winner) side of their momentum portfolio. Moreover, Boguth et al. (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) are biased. Lettau et al. (2014) propose a downside risk capital asset pricing model (DR-CAPM) which they find explains the cross section of returns in many asset classes better.

10

where $R^e_{\mathrm{MKT},t}$ is the market portfolio returns in excess of risk free return for month t. We note that α_p, the intercept of the regression, is no longer a measure of the strategy's abnormal return, because the option payoff—$\max(R^e_{\mathrm{MKT},t}, 0)$—is not an excess return. We return to this issue and estimate the abnormal return of the strategy in Section 5. For the moment, we concentrate on the time-variation in β^+, which is a measure of the exposure of the portfolio p to the payoff on a one-month call option on the stock market or, equivalently, a measure of the convexity in the relationship between the market return and the momentum strategy return.

To examine this time-variation, we partition the months in our sample into three groups on the basis of three state variables: the cumulative market return during the 36 month preceding the portfolio formation month; the realized volatility of daily market returns over the previous 12 months; and the breakpoints of the loser portfolio – i.e., the return over the (12,2) measurement period of the stock at the 10th percentile. Based on each of these state variables, we partition our sample of 1044 months into 'High', 'Medium' and 'Low' groups. The High (Low) group is the set of months when the state variable is in the top (bottom) 20th percentile at the start of that month. The 'Medium' group contains the remaining months (i.e., the middle 60%). We present the results from sorting on the basis of the past 36-month market return in Table 3; the results from sorting on the other two state variables are presented in Table 14 of the Online Appendix.[17]

Panel A presents the estimates of equation (1) for the momentum strategy returns (R_{MOM}), and for the returns of the winner and loser portfolio in excess of the risk free rate (R^e_{WIN} and R^e_{LOS}). First, note that the estimated β^+, the exposure to the market call payoff is significant only when the past 36-month market returns are in 'Low' group: consistent with the leverage hypothesis, the past-loser has a positive exposure to the market

---

[17]Results are similar when we group based on other variables that capture market conditions: the cumulative market return during the 12 month preceding the portfolio formation month; the realized volatility of daily market returns over the previous 6 months and the ratio of the book value of debt to the market value of equity (BD/MV) of the loser stock portfolio.

option payoff of 0.72 ($t = 3.60$). That is, it behaves like a call option on the market. The MOM portfolio, which is short the past-losers, thus has a significantly negative β^+.[18] In the 'Medium' and 'High' groups, the β^+ of the MOM returns and of the long- and short-sides are smaller in absolute value and are not statistically significantly negative.[19] In the 'Low' state, the $Adj.R^2$ is 48% for MOM returns, as compared to 6% in both the 'Medium' and 'High' states, a result of both the higher β^0 and β^+ in the Low state.

Panel C shows that large MOM losses (crashes) are clustered in months when the option-like feature of β^+ is accentuated; 11 out of 13 momentum losses occur during months in the 'Low' group. Table 14 shows the results when we group samples by the other state variables: i.e., realized volatility of market over the past 12 months or return breakpoints for stocks to enter the loser portfolio.

The evidence in Panel D suggests that the large negative skewness of the momentum strategy return distribution is mostly due to the embedded written call option on the market. In the 'Low' group of Panel A, the skewness of the momentum strategy returns is -2.33, but after we control for the non-linear exposure to the market through equation (1), the skewness of residual drops to -0.48. In 'Medium' and 'High' group, the negative skewness of momentum strategy returns is not that strong and it is not significantly reduced after controlling for the embedded written call option on the market. This is consistent with the results in Panel A; β^+ is not significantly different from zero in the other two groups. The results reported in Table 14 of the Online Appendix are consistent with the results presented here: the large negative skew in momentum returns is due to the embedded written call option that gets accentuated by market conditions.

The above results suggest that the embedded written call option on the market is the

---

[18]The t-statistics are computed using the heteroskedasticity-consistent covariance estimator of White (1980).

[19]We note that β^+ of winner and loser portfolios exhibit interesting patterns: It is negative and significant for winner stocks in 'Low' group. It is negative and statistically significant for loser stocks in the 'High' group. Understanding why we see these patterns is left for future research.

key driver of momentum crashes, and that this optionality is a result of the high leverage of the past-loser firms. However this leverage will not always be apparent in the *financial* leverage of the past-loser portfolio. For example, it is likely that the *operating* leverage of many of the firms that earned low returns in the post-March 2000 collapse of the tech sector was quite high, even though these firms' financial leverage was insignificant. This is consistent with the evidence that the financial leverage of the loser portfolio was low during two episodes of large momentum losses in 2001:01 and 2002:12.[20] However, as can be seen from Table 4, the optionality is large when we estimate the augmented market model return generating process for momentum returns given by equation (1) for the 36 monthly returns from 2000:01-2002:12—although it is not statistically significant due to the small sample size.

In the next section, we model the option-like relationship between the market and the momentum portfolio, with the goal of employing this model to forecast momentum crashes. The evidence above suggests that a model based on Merton (1974), using debt and equity values would not capture these periods. Alternatively, we could form a model with a functional form relating the state-variables explored above (past-market returns, market volatility, etc.) to the convexity of momentum returns. This, however, requires choosing the length of time window over which these state-variables are measured, and that necessarily has to be arbitrary. Given these difficulties, we instead posit a two-state model, with "calm" and "turbulent" states. When the economy is in the turbulent state, the option like feature of momentum returns is accentuated, and momentum crashes are more likely. This naturally leads us to the two-state hidden Markov model (HMM) for identifying time periods when momentum crashes are more likely, which we explore in the next Section.

---

[20]Refer to Table 2. In 2001:01 (2002:12), the momentum strategy loses -41.97% (-20.40%) and the financial leverage (BD/MV) of loser portfolio was 0.68 (2.32). The average of financial leverage over all available data from 1964 is 3.97.

3 The Two-State Hidden Market Model

In this section we develop a two-state hidden Markov model (HMM) in which a single state variable summarizes the market conditions. The "turbulent" state is characterized by higher return volatilities and by more convexity in the market-momentum return relationship. We then show how the HMM allows *ex-ante* estimation of the probability that the hidden state is *calm* or *turbulent* based on the history of momentum and market returns.

3.1 A Hidden Markov Model of Market and Momentum Returns

Let S_t denote the unobserved underlying state of the economy at time t, which is either "*calm*" (C) or "*turbulent*" (T) in our setting. Our specification for the return generating process of the momentum strategy is as follows:

$$R_{\text{MOM},t} = \alpha(S_t) + \beta^0(S_t)R_{\text{MKT},t}^e + \beta^+(S_t)\max\left(R_{\text{MKT},t}^e, 0\right) + \sigma_{\text{MOM}}(S_t)\,\varepsilon_{\text{MOM},t}, \quad (2)$$

where $\varepsilon_{\text{MOM},t}$ is an *i.i.d* random process with zero mean and unit variance. Equation (2) is similar to equation (1). However, the option-like feature, $\beta^+(S_t)$, the sensitivity of momentum strategy return to the market return, $\beta^0(S_t)$, and the volatility of momentum specific shock, $\sigma_{\text{MOM}}(S_t)$, all differ across the unobserved turbulent and calm states of the economy. We also let the intercept, $\alpha(S_t)$, vary across the two hidden states of the economy. We assume that the return generating process of the market return in excess of risk free rate is given by:

$$R_{\text{MKT},t}^e = \mu(S_t) + \sigma_{\text{MKT}}(S_t)\varepsilon_{\text{MKT},t}, \quad (3)$$

where $\varepsilon_{\text{MKT},t}$ is an *i.i.d* random process with zero mean and unit variance. That is, $\mu(S_t)$ and $\sigma_{\text{MKT}}(S_t)$ represent the state dependent mean and volatility of the market excess return.

Finally, we assume that the transition of the economy from one hidden state to another

is Markovian, with the transition probability matrix as given below:

$$
\Pi = \left[\begin{array}{cc} \Pr(S_t = C | S_{t-1} = C) & \Pr(S_t = T | S_{t-1} = C) \\ \Pr(S_t = C | S_{t-1} = T) & \Pr(S_t = T | S_{t-1} = T) \end{array} \right], \tag{4}
$$

where S_t, the unobservable random state at time t which, in our setting, is either $Calm(C)$ or $Turbulent(T)$ and $\Pr(S_t = s_t | S_{t-1} = s_{t-1})$ denotes the probability of transitioning from state s_{t-1} at time $t-1$ to state s_t at time t.[21]

3.2   Maximum Likelihood Estimation

We now estimate the parameters of the hidden Markov model in equations (2), (3), and (4), which we summarize with the 14-element parameter vector θ^0:

$$
\theta^0 = \left\{ \begin{array}{c} \alpha\left(C\right), \beta^0\left(C\right), \beta^+\left(C\right), \sigma_{\text{MOM}}\left(C\right), \\ \alpha\left(T\right), \beta^0\left(T\right), \beta^+\left(T\right), \sigma_{\text{MOM}}\left(T\right), \\ \mu\left(C\right), \sigma_{\text{MKT}}\left(C\right), \mu\left(T\right), \sigma_{\text{MKT}}\left(T\right), \\ \Pr\left(S_t = C | S_t = C\right), \Pr\left(S_t = \text{T} | S_t = \text{T}\right) \end{array} \right\}. \tag{5}
$$

The observable variables are the time series of excess returns on the momentum portfolio and on the market, which we summarize in the vector $\mathbf{R}_t$:

$$
\mathbf{R}_t = \left(R_{\text{MOM},t}, R^e_{\text{MKT},t} \right)'.
$$

We let $\mathbf{r}_t$ denote the realized value of $\mathbf{R}_t$.

We follow Hamilton (1989) and estimate the HMM parameters by maximizing the log likelihood of the sample, given distributional assumptions for $\varepsilon_{\text{MOM},t}$ and $\varepsilon_{\text{MKT},t}$, in (2) and

---

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

(3). As shown in Appendix A, when the likelihood is misspecified, the ML estimator of θ^0 can be inconsistent. Hence, we choose the distribution of $\varepsilon_{\mathrm{MOM},t}$ and $\varepsilon_{\mathrm{MKT},t}$ so that the unconditional variance, skewness, and kurtosis of momentum and market excess returns implied by our HMM specification are closer to their sample analogues. As we discuss later in more detail, while the momentum returns $R_{\mathrm{MOM},t}$ are highly skewed and leptokurtic, the momentum return residuals ($\varepsilon_{\mathrm{MOM},t}$) appear normally distributed. Interestingly, the market return residual ($\varepsilon_{\mathrm{MKT},t}$) is non-normal—it is better characterized as Student's t-distributed with d.f.=5.

Let $\widehat{\theta}_{\mathrm{ML}}$ denote the vector of HMM parameters that maximizes the log likelihood function of the sample given by:

$$\mathcal{L} = \sum_{t=1}^{T} \log\left(\Pr\left(\mathbf{r}_t | \mathcal{F}_{t-1}\right)\right), \tag{6}$$

where $\mathcal{F}_{t-1}$ denotes the agent's time $t-1$ information set (*i.e.* all market and momentum excess returns up through time $t-1$).

Given the hidden-state process that governs returns, the time-t element of this equation—the likelihood of observing $\mathbf{r}_t$—is:

$$\Pr\left(\mathbf{r}_t | \mathcal{F}_{t-1}\right) = \sum_{s_t \in \{C,T\}} \Pr\left(\mathbf{r}_t, S_t = s_t | \mathcal{F}_{t-1}\right), \tag{7}$$

where the summation is over the two possible values of the unobservable state variable S_t. The joint likelihood inside the summation can be written as:

$$\Pr\left(\mathbf{r}_t, S_t = s_t | \mathcal{F}_{t-1}\right) = \Pr\left(\mathbf{r}_t | S_t = s_t, \mathcal{F}_{t-1}\right) \Pr\left(S_t = s_t | \mathcal{F}_{t-1}\right)$$

$$= \Pr\left(\mathbf{r}_t | S_t = s_t\right) \Pr\left(S_t = s_t | \mathcal{F}_{t-1}\right). \tag{8}$$

The first term of equation (8) is the state dependent likelihood of $\mathbf{r}_t$ which can be computed,

16

given distributional assumptions for $\varepsilon_{\text{MOM},t}$ and $\varepsilon_{\text{MKT},t}$ in (2) and (3) as:

$$\Pr\left(\mathbf{r}_t | S_t = s_t\right) = \Pr\left(\varepsilon_{\text{MOM},t} | S_t = s_t\right) \cdot \Pr\left(\varepsilon_{\text{MKT},t} | S_t = s_t\right)$$

where

$$\varepsilon_{\text{MOM},t} = \frac{1}{\sigma_{\text{MOM}}\left(s_t\right)} \left(r_{\text{MOM},t} - \alpha\left(s_t\right) - \beta^0\left(s_t\right) r^e_{\text{MKT},t} - \beta^+\left(s_t\right) \max\left(r^e_{\text{MKT},t}, 0\right)\right)$$

$$\varepsilon_{\text{MKT},t} = \frac{1}{\sigma_{\text{MKT}}\left(s_t\right)} \left(r^e_{\text{MKT},t} - \mu\left(s_t\right)\right).$$

The second term of equation (8) can be written as a function of the time $t-1$ state probabilities as:

$$
\begin{aligned}
\Pr\left(S_t = s_t | \mathcal{F}_{t-1}\right) &= \sum_{s_{t-1} \in \{C,T\}} \Pr\left(S_t = s_t, S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}\right) \\
&= \sum_{s_{t-1} \in \{C,T\}} \Pr\left(S_t = s_t | S_{t-1} = s_{t-1}, \mathcal{F}_{t-1}\right) \Pr\left(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}\right) \\
&= \sum_{s_{t-1} \in \{C,T\}} \Pr\left(S_t = s_t | S_{t-1} = s_{t-1}\right) \Pr\left(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}\right), \quad (9)
\end{aligned}
$$

where third equality holds since the transition probabilities depend only on the hidden state. We can compute the expression on the left hand side of equation (9) using the elements of the transition matrix, $\Pr\left(S_t = s_t | S_{t-1} = s_{t-1}\right)$. The right hand side of equation (9)—the conditional state probability $\Pr\left(S_{t-1} = s_{t-1} | \mathcal{F}_{t-1}\right)$—comes from Bayes' rule:

$$
\begin{aligned}
\Pr\left(S_t = s_t | \mathcal{F}_t\right) &= \Pr\left(S_t = s_t | \mathbf{r}_t, \mathcal{F}_{t-1}\right) \\
&= \frac{\Pr\left(\mathbf{r}_t, S_t = s_t | \mathcal{F}_{t-1}\right)}{\Pr\left(\mathbf{r}_t | \mathcal{F}_{t-1}\right)}. \quad (10)
\end{aligned}
$$

where the numerator and denominator of equation (10) come from equations (8) and (7), respectively.

17

Thus, given time 0 state probabilities, we can calculate the conditional state probabilities for all $t \in \{1, 2, \cdots, T\}$. In our estimation, we set $\Pr(S_0 = s_0 | \mathcal{F}_0)$ to their corresponding steady state values implied by the transition matrix.[22] Table 5 gives the Maximum Likelihood parameter estimates and standard errors of the hidden Markov model parameter vector in equation (5) with our assumption that the momentum return residual ($\varepsilon_{\mathrm{MOM},t}$) is drawn from standard normal distribution and the market returns residual is drawn from a Student's t-distribution with d.f.=5, which will be justified later.

The parameters in Table 5 suggest that HMM does a good job of picking out two distinct states: Notice that β^+, while still negative in the calm state, is more than twice as large in the turbulent state. Similarly, the estimated momentum and market return volatilities, $\sigma_{\mathrm{MOM}}(S_t)$ and $\sigma_{\mathrm{MKT}}(S_t)$, are more than twice as large in the turbulent state. We see also that the calm state is more persistent than the turbulent, at least based on point estimates.

An implication of the large $\beta^+(T)$ is that MOM's response to up moves in the market is considerably more negative than the response to down-moves in the market. In the turbulent state, MOM's up market beta is -1.45 (=-0.32-1.14), but its down market beta is only -0.32. The combination of this feature and the higher volatilities means that the left tail risk is high when the hidden state is turbulent.

One rather striking feature of the numbers in Table 5 is the large differences in the market parameters across the two states. For the calm state, the point estimates of the annualized expected excess return and volatility of the market are, respectively, 13.3%/year and 14.0%, giving an annualized Sharpe-ratio is 0.95. In contrast, the corresponding estimated parameters for the turbulent state are -4.6%/year and 29.0%. We caution the reader that the hidden state is not observed, so these returns are not directly achievable. We also note that these results are consistent with prior evidence on the inverse relationship between market volatility and market Sharpe Ratios (Glosten et al., 1993; Breen et al., 1989; Moreira and

---

[22]The vector of steady state probabilities is given by the eigenvector of the transition matrix given in equation (4).

Muir, 2015).

A natural question that arises is whether our HMM specification is consistent with the highly non-normal momentum return distribution in our sample. We therefore compare the unconditional sample moments of momentum strategy returns and market excess returns implied by the HMM return generating process with their sample counterparts. For this purpose, we consider the following distributions for ($\varepsilon_{\mathrm{MOM},t}$, $\varepsilon_{\mathrm{MKT},t}$): (Normal, Student's t), (Normal, Normal), (Student's t, Normal), and (Student's t, Student's t). For each of these pairs of distributions, we estimate our HMM model, generate a 1044 month-long time series of momentum strategy and market excess returns using Monte Carlo simulation and obtain their first four moments. We then repeat this exercise 10,000 times to obtain the distribution of the first four momentums implied by the HMM specification. Table 6 summarizes the distribution of the first four moments of the momentum strategy returns and market excess returns obtained in this way for the four pairs of distributions.

Panel A of Table 6 gives the result for our baseline specification of normal ($\varepsilon_{\mathrm{MOM},t}$) and Student's t ($\varepsilon_{\mathrm{MKT},t}$), which is the only case where all of the first four realized moments of momentum strategy returns, over our sample period 1044 months (1927:01-2013:12), fall inside the 95% confidence interval for the HMM-implied moments obtained by simulation.[23] These findings are consistent with the hypothesis that the significant left skewed and leptokurtic sample momentum strategy returns are due to the non-linear exposure to market returns. In contrast, when we use normal distributions for $\varepsilon_{\mathrm{MKT},t}$, the sample skewness of momentum strategy returns lies outside of the 99 (95) % confidence interval of our HMM-implied moments, as can be seen from Panel B (C) of Table 6. Furthermore, the sample kurtosis of market excess returns exceeds the 99.5th percentile value of its HMM-implied distribution. If we use Student's t-distributions for both $\varepsilon_{\mathrm{MOM},t}$ and $\varepsilon_{\mathrm{MKT},t}$, the realized skewness of mo-

---

[23]If we perform the non-parametric test by Kolmogorov-Smirnov on the similarity between the empirical CDFs of realized momentum returns and simulated momentum returns, any of the four distributional assumptions is not rejected with 5% significance level due to the low power of the test. Hence, we examine the distribution for each of the first four moments.

mentum returns lies outside of the 95% confidence intervals of our HMM-implied moments and the confidence intervals for kurtosis becomes much wider. When we compare the distribution of kurtosis in Panel A with that in Panel D, the 95% confidence intervals of kurtosis of momentum strategy returns is (6.87, 36.30) when $\varepsilon_{\mathrm{MOM},t}$ have a Student's t-distribution – much wider than the corresponding 95% confidence intervals of (5.63, 29.79) when $\varepsilon_{\mathrm{MOM},t}$ has a normal distribution. Given these finding, we assume that $\varepsilon_{\mathrm{MOM},t}$ is drawn from a normal distribution and $\varepsilon_{\mathrm{MKT},t}$ is drawn from a Student's t-distribution with 5 degrees of freedom.

We now proceed to examine the extent to which the estimated state probabilities can forecast the momentum "crashes" observed in our sample.

4   Using the HMM to Predict Momentum Crashes

In this section, we examine the predictability of momentum crashes based on the estimated probability of the economy being in the hidden turbulent state in a given month, $\Pr(S_t = \mathrm{T}|\mathcal{F}_{t-1})$. It is evident from Table 5 that when the hidden state is turbulent, the written call option-like features of momentum strategy returns become accentuated, and both the momentum strategy and market excess returns become more volatile. Hence, we should expect that the frequency with which extreme momentum strategy losses occur should increase with $\Pr(S_t = \mathrm{T}|\mathcal{F}_{t-1})$.

Figure 2 presents scatter plots of realized momentum strategy returns on the vertical axis against $\Pr(S_t = \mathrm{T}|\mathcal{F}_{t-1})$, the estimated probability that the hidden state is turbulent, on the horizontal axis. Momentum strategy losses exceeding 20% are in red and momentum strategy gains exceeding 20% are in green. Panel A is based on in-sample estimates using all 1044 months of data during 1927:01-2013:12. Consistent with results in the preceding section, the large losses, highlighted in red, occur only when the estimated turbulent state

probability is high. The large gains (the green dots) are fairly evenly distributed across the different state probabilities.

The analysis reflected in Panel A is in-sample, meaning that the full-sample parameters (i.e., those presented in Table 5) are used to estimate the state probability at each point in time. In Panel B, the turbulent state probability is estimated fully out-of-sample; the parameters are estimated by the same QML procedure, but only up through the month prior to portfolio formation. Here the sample is 1980:09-2013:12, giving us a sufficiently large period over which to estimate the parameters. To further challenge the HMM estimation, we estimate the HMM parameters using only data from the slightly less volatile period following 1937:01. In Panel B, just as in Panel A, there is strong association between momentum crashes worse than -20% (red dots) and high values of the (out-of-sample) estimated turbulent state probability. In contrast, large momentum gains more than 20% (green dots) are dispersed more evenly across high and low values of the estimated state probability.

Table 7 presents the number of large negative and large positive momentum strategy returns during months when $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$ is above a certain threshold. Notice that all thirteen momentum crashes, defined as losses exceeding 20%, happen when the $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$ is more than 80%. However, only eight out of twelve momentum gains exceeding 20% are found when the $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$ is more than 80%, and three out of those large gains happen when the $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$ is less than 30%.

Most quantitative fund managers operate with mandates that impose limits on their portfolios' return-volatilities. Barroso and Santa-Clara (2015) demonstrate the benefit of such mandates: when exposure to the momentum strategy is varied over time to keep its volatility constant the Sharpe ratio significantly improves. A natural question that arises is whether managing the volatility of the portfolio to be within a targeted range is the best way to manage the portfolio's exposure to left tail risk. We add to this literature by focusing on tail risk, i.e., the probability of very large losses. As we saw before, left tail risk is related

to left skewness of returns, and there are no a priori reasons to believe that changes in left skewness move in lock step with changes in the volatility of momentum strategy returns. We therefore let the data speak, by comparing the performance of two tail risk measures: the volatility of momentum strategy returns (measured either by realized volatility or by GARCH) and the probability of the economy being in a turbulent state computed based on the estimated HMM parameters in predicting momentum crashes.

Table 8 compares the number of false positives in predicting momentum crashes across different tail risk measures. The number of false positives of a given tail risk measure is computed as follows. Suppose we classify months in which momentum strategy returns lost more than a threshold X. Let Y denote the lowest value attained by a given tail risk measure during those momentum crash months. For example, consider all months during which the momentum strategy lost more than 20% ($X=20\%$). Among those months, the lowest value, attained by the tail risk measure of $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$, is 84% ($Y=84\%$). During months when the tail risk measure is above the threshold level of Y, we count the number of months when momentum crashes did not occur and we denote it as the number of false positives. Clearly, the tail risk measure that has the least number of false positives is preferable. Table 8 gives the number of false positives for different tail risk measures and different values of threshold $X=10\%$, 20%, 30%, 40%.

In Panel A, we use $\Pr\left(S_t = \text{T}|\mathcal{F}_{t-1}\right)$ as a tail risk measure. The results in Panel A-1 are from our original HMM model specified in (2), (3) and (4). To emphasize the importance of the option-like feature $\beta^+(S_t)$ in (2), we impose the restriction $\beta^+(S_t) = 0$ and report the associated results in Panel A-2.

In Panel B, we use various estimates of the volatility of momentum strategy returns as tail risk measures. Specifically, we estimate the volatility of the momentum strategy returns using GARCH (1,1), and realized volatility of daily momentum strategy returns over the previous 3, 6, 12, and 36 months. In Panel C, we use the volatility of the market return

estimated using GARCH(1,1) which can be viewed as realized volatility using all past returns and realized volatility of the daily market return during the preceding 3, 6, 12, and 36 months as tail risk measures. In Panel D, we use the market return during the preceding 3, 6, 12 and 36 month windows as tail risk measures.

When $X=20\%$, we find that the number of false positives in Panel A-1 is always smaller than other cases in Panel B, C and D. For example, in our 1930:01-2013:12 sample,[24] we find 114 false positives when we use the tail risk measure based on our main specification of HMM. In contrast, if we use the realized volatility of daily momentum strategy returns over the previous six months,[25] the number of false positives increases to 187 months. The result of Panel A-2 shows the necessity of the option features in our HMM specification. If we impose that $\beta^+(S_t) = 0$ while estimating our HMM model, the performance becomes worse. The number false positives increases from 114 to 150.

This establishes the link between the tail risk of momentum strategy returns and the probability of the economy being in the hidden turbulent state. In the next section we examine how the alpha of the momentum strategy return varies over time, as the probability of the economy being in the turbulent state changes.

5 Momentum's Option Adjusted Alpha

We have shown that the two-state HMM effectively picks out changes in the market environment that lead to dramatic shifts in the distribution of market and momentum returns. Moreover, even when estimated out-of-sample, the HMM does a far more effective job of forecasting momentum tail events or "crashes" than alternative methods.

These results raise the question of how the alpha of the momentum strategy varies over

---

[24]Since we utilize momentum returns over the previous 36 months to construct risk measures, the sample period becomes shorter.

[25]Barroso and Santa-Clara (2015) used this measure to imposing the volatility target of the momentum strategy.

time with changes in market conditions. While not the focus of our paper, in this section we briefly examine this question, based on the estimated HMM model from Section 3. We calculate the alpha from the perspective of an investor who can freely invest in the risk free asset, the market index portfolio, and at-the-money call options on the market index portfolio without any frictions, but whose pricing kernel is otherwise uncorrelated with innovations in the momentum strategy. Given this assumption, our valuation requires the prices of traded options on the market portfolio, which we proxy with one month, at-the-money index options on the S&P 500.

Specifically, we assume that how the investor values payoffs on risky assets has the following stochastic discount factor representation.[26] Let M_t denote the stochastic discount factor, and $\mathcal{F}_{t-1}$ the investor's information set at time $t-1$. Since the investor has frictionless access to the risk free asset, the market portfolio, and call options on the market portfolio, the followings relations hold:

$$
\begin{aligned}
1 &= \mathbb{E}\left[M_t(1 + R_{\mathrm{f},t})|\mathcal{F}_{t-1}\right] \\
0 &= \mathbb{E}\left[M_t R_{\mathrm{MKT},t}^e|\mathcal{F}_{t-1}\right] \\
V_{c,t-1} &= \mathbb{E}\left[M_t \max\left(R_{\mathrm{MKT},t}^e, 0\right)|\mathcal{F}_{t-1}\right],
\end{aligned}
$$

where $R_{\mathrm{f},t}$ is the risk free rate from $t-1$ to t and $V_{c,t-1}$ is the market price of the call option which pays $\max\left(R_{\mathrm{MKT},t}^e, 0\right)$ at the end of time t.

Regress M_t, which is based on a constant, the market excess return, and the payoff of a call option on the market based on the information set $\mathcal{F}_{t-1}$. Let $\widetilde{M_t}$ be the fitted part of M_t and $\widetilde{e}_t$ be the residual in that conditional regression. Then we can write M_t as follows:

$$
M_t = \widetilde{M_t} + \widetilde{e}_t \tag{11}
$$

---

[26]In our derivations, we follow the framework in Hansen and Jagannathan (1991) and Glosten and Jagannathan (1994).

where

$$\widetilde{M}_t = \lambda_{0,t-1} + \lambda_{1,t-1} R^e_{\text{MKT},t} + \lambda_{2,t-1} \left(R^e_{\text{MKT},t}, 0 \right) \tag{12}$$

$$\mathbb{E}\left[\widetilde{e}_t | \mathcal{F}_{t-1} \right] = \mathbb{E}\left[R^e_{\text{MKT},t} \widetilde{e}_t | \mathcal{F}_{t-1} \right] = \mathbb{E}\left[\max \left(R^e_{\text{MKT},t}, 0 \right) \widetilde{e}_t | \mathcal{F}_{t-1} \right] = 0. \tag{13}$$

The residual $\widetilde{e}_t$ represents the risk that the investor cares about that is not an affine function of the risk free return, market excess return, and the payoff of the call option on the market excess return. $\widetilde{M}_t$ is the dynamic analogue of the stochastic discount factor implied by the downside risk capital asset pricing model proposed by Lettau et al. (2014)

In a similar manner, regress the momentum strategy return on a constant, the market excess return, and the call option payoff on the market given the information set $\mathcal{F}_{t-1}$. Recall that when the hidden state S_t is turbulent, which occurs with the probability of $\Pr\left(S_t = T | \mathcal{F}_{t-1} \right)$, the momentum strategy return and market excess return generating processes are given by equation (2), where S_t is either calm or turbulent, and where $\varepsilon_{\text{MOM},t}$ and $\varepsilon_{\text{MKT},t}$ are assumed to be drawn from standard normal and Student's t-distributions, respectively.

We consider the following conditional regression given the information set $\mathcal{F}_{t-1}$ that includes the risk free return and the price of the call option on the market:

$$R_{\text{MOM},t} = \alpha_{t-1} + \beta^0_{t-1} R^e_{\text{MKT},t} + \beta^+_{t-1} \max \left(R^e_{\text{MKT},t}, 0 \right) + \epsilon_{\text{MOM},t}, \tag{14}$$

where

$$\mathbb{E}\left[\epsilon_{\text{MOM},t} | \mathcal{F}_{t-1} \right] = \mathbb{E}\left[\epsilon_{\text{MOM},t} R^e_{\text{MKT},t} | \mathcal{F}_{t-1} \right] = \mathbb{E}\left[\epsilon_{\text{MOM},t} \max \left(R^e_{\text{MKT},t}, 0 \right) | \mathcal{F}_{t-1} \right] = 0.$$

Specifically, the vector of regression coefficients $\begin{bmatrix} \alpha_{t-1} & \beta_{t-1}^0 & \beta_{t-1}^+ \end{bmatrix}'$ is determined as

$$\begin{bmatrix} \alpha_{t-1} & \beta_{t-1}^0 & \beta_{t-1}^+ \end{bmatrix}' = (\mathbb{E}\left[\mathbf{x}_t\mathbf{x}_t'|\mathcal{F}_{t-1}\right])^{-1}\mathbb{E}\left[\mathbf{x}_t R_{\text{MOM},t}|\mathcal{F}_{t-1}\right]$$

where $\mathbf{x}_t = \begin{bmatrix} 1 & R_{\text{MKT},t}^e & \max\left(R_{\text{MKT},t}^e, 0\right) \end{bmatrix}'$, and

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{x}_t\mathbf{x}_t'|\mathcal{F}_{t-1}\right] &= \Pr\left(S_t = C|\mathcal{F}_{t-1}\right)\mathbb{E}\left[\mathbf{x}_t\mathbf{x}_t'|S_t = C\right] + \Pr\left(S_t = T|\mathcal{F}_{t-1}\right)\mathbb{E}\left[\mathbf{x}_t\mathbf{x}_t'|S_t = T\right] \\
\mathbb{E}\left[\mathbf{x}_t R_{\text{MOM},t}|\mathcal{F}_{t-1}\right] &= \Pr\left(S_t = C|\mathcal{F}_{t-1}\right)\mathbb{E}\left[\mathbf{x}_t R_{\text{MOM},t}|S_t = C\right] \\
&\quad + \Pr\left(S_t = T|\mathcal{F}_{t-1}\right)\mathbb{E}\left[\mathbf{x}_t R_{\text{MOM},t}|S_t = T\right].
\end{aligned}
$$

Furthermore, the regression equation of (14) can be expressed in terms of excess returns as follows:

$$R_{\text{MOM},t} = \alpha_{t-1}^* + \beta_{t-1}^0 R_{\text{MKT},t}^e + \beta_{t-1}^+ V_{c,t-1}\left(\frac{\max\left(R_{\text{MKT},t}^e, 0\right)}{V_{c,t-1}} - (1 + R_{f,t})\right) + \epsilon_{\text{MOM},t}, \quad (15)$$

where the quantity in parenthesis is the excess return on one-period call option on the market.[27]

$$\alpha_{t-1}^* = \alpha_{t-1} + (1 + R_{f,t})\beta_{t-1}^+ V_{c,t-1}. \quad (16)$$

We denote α_{t-1}^* as the option adjusted alpha of the momentum strategy return. When the following assumption holds, $\frac{\alpha_{t-1}^*}{1+R_{f,t}}$ gives the marginal value of the momentum strategy return from the perspective of the marginal investor.

Assumption 1. $\mathbb{E}\left[\widetilde{e}_t\epsilon_{\text{MOM},t}|\mathcal{F}_{t-1}\right] = 0$ *where* $\widetilde{e}_t$ *and* $\epsilon_{\text{MOM},t}$ *are given in equations (11) and (15), respectively.*

With Assumption 1, the following proposition holds.

---

[27]The strike price of the option is the level of the market index times $(1 + R_{f,t})$, which means that the option will be at-the-money at expiration if $R_{\text{MKT},t}^e = 0$.

Proposition 1. *The value of momentum strategy return to the investor whose stochastic discount factor is M_t, is $\frac{\alpha^*_{t-1}}{1+R_{f,t}}$.*

Proof.

$$
\mathbb{E}\left[M_t R_{\mathrm{MOM},t} | \mathcal{F}_{t-1}\right]
$$

$$
= \alpha^*_{t-1}\mathbb{E}\left[M_t|\mathcal{F}_{t-1}\right] + \beta^0_{t-1}\mathbb{E}\left[M_t R^e_{\mathrm{MKT},t}|\mathcal{F}_{t-1}\right]
$$

$$
+ \beta^+_{t-1}V_{c,t-1}\left(\frac{\mathbb{E}\left[M_t \max\left(R^e_{\mathrm{MKT},t}, 0\right)|\mathcal{F}_{t-1}\right]}{V_{c,t-1}} - (1+R_{f,t})\mathbb{E}\left[M_t|\mathcal{F}_{t-1}\right]\right)
$$

$$
+ \mathbb{E}\left[M_t \epsilon_{\mathrm{MOM},t}|\mathcal{F}_{t-1}\right]
$$

$$
= \frac{\alpha^*_{t-1}}{1+R_{f,t}} + \mathbb{E}\left[\left(\widetilde{M}_t + \widetilde{e}_t\right)\epsilon_{\mathrm{MOM},t}|\mathcal{F}_{t-1}\right] = \frac{\alpha^*_{t-1}}{1+R_{f,t}} + \mathbb{E}\left[\widetilde{e}_t \epsilon_{\mathrm{MOM},t}|\mathcal{F}_{t-1}\right]
$$

$$
= \frac{\alpha^*_{t-1}}{1+R_{f,t}},
$$

where the first equality follows from equation (15). The second equality follows from the assumption that the investor, whose stochastic discount factor is M_t, agrees with the market prices of the risk free asset, market excess return, the call option payoff and the decomposition in equation (11). The third equality follows from equation (12) and the properties of the conditional regression residual $\epsilon_{\mathrm{MOM},t}$. The last equality follows from Assumption 1. $\qquad\square$

In what follows we compute the time series of the estimated option adjusted alpha, α^*, in (16) based on the time series of risk free returns and the prices of call options. We then assess the validity of Assumption 1 by examining whether the residual in the equation (15) is uncorrelated with various risk factors proposed in the literature. Figure 3 plots the time series of α^* calculated based on the estimated HMM model for the sample period 1996:01-2013:12. Notice that the sample average of the α^*_{t-1}'s is 1.17%/month, which is significantly positive. However, α^*_{t-1} is negative during 1998:09-1998:10 (Russian crisis), 2002:09-2002:10 (dot-com bubble bursts), 2008:10-2008:12 and 2009:02-2009:04 (financial crisis) – months when option prices were high and the market was more likely to be in the hidden turbulent

state.

We compute the confidence intervals for the estimated option adjusted alphas as follows. First, we simulate 10,000 sets of parameters from the asymptotic distributions obtained from QML estimator, reported in Table 5. Then, for each set of parameters, we estimate the probability for the hidden state being turbulent based on the realized market excess returns and momentum strategy returns in our sample period 1996:01-2013:12 . With the simulated parameters, the estimated probabilities, and the time series of risk-free returns and call option prices, we construct the time series of α^*'s for the period 1996:01-2013:12 as described earlier. Finally, for each month, we find the 95% confidence intervals of α^* by choosing the top and bottom 2.5% quantiles from the simulated 10,000 α^* in each month.

In Figure 3, we plot the time series of estimated α_{t-1}^* along with the corresponding 95% confidence intervals. In 167 of the 216 months in the sample period 1996:01-2013:12, the option-adjusted alpha is significantly positive. While the option adjusted alpha is negative during 10 months, only during three months – both occur during the recent financial crisis period 2008:11-2008:12 and 2009:03 – are they statistically significantly different from zero.

To assess the reasonableness of Assumption 1, we construct the time series of the residuals, $\epsilon_{\text{MOM},t}$ in equation (15), based on the estimated parameter values as follows.

$$\epsilon_{\text{MOM},t} = R_{\text{MOM},t} - \alpha_{t-1}^* - \beta_{t-1}^0 R_{\text{MKT},t}^e - \beta_{t-1}^+ V_{c,t-1}\left(\frac{\max\left(R_{\text{MKT},t}^e, 0\right)}{V_{c,t-1}} - (1 + R_{\text{f},t})\right).$$

We regress the residual on commonly used economy wide risk factors in the literature: the three factors of market excess returns (MKT), small minus big size (SMB), high minus low book to market (HML) in Fama and French (1993); robust minus weak (RMW) and conservative minus aggressive (CMA) factor in Fama and French (2015); investment to assets (I/A) and return on equity (ROE) factor in Hou et al. (2015); quality minus junk (QMJ) factor in Asness et al. (2014); liquidity risk factor (LIQ) in Pastor and Stambaugh (2003);

funding liquidity risk factor (FLS) in Chen and Lu (2015); betting against beta (BAB) risk factor in Frazzini and Pedersen (2014); changes in 3-Month LIBOR (LIBOR), Term Spread (the yield spread between the 10-year treasury bond and 3-month T-bill, TERM), Credit Spread (the yield spread between Moody's BAA bond and AAA bond, CREDIT), and TED Spread (the yield spread between the 3-month LIBOR and 3-month T-bill, TED); and returns of variance swap (VAR-SWAP) across different horizons (Dew-Becker et al., 2015); and the changes in VIX as well as the changes in left jump variations (LJV) embedded in option prices measured by Bollerslev et al. (2015).[28] Specifically, we estimate the regression equation

$$\epsilon_{\mathrm{MOM},t} = \mathrm{intercept} + \mathrm{coeff} \times \mathrm{systematic\ factor}_t + e_t$$

and report coeff (t-stat) and R^2 in Table 9. Except for HML and ROE factors, we do not find any significant correlation between the residual we computed and systematic risk factors.[29] These findings suggest the need for using dynamic versions of the Fama and French (1993) three-factor model and the Hou et al. (2015) four-factor model to value momentum strategy returns in any given month based on all information available at that point in time. We leave such extensions for future research.

---

[28]We obtain MKT, SMB, HML, CMA and RMW from Ken French's data library: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html; QMJ and BAB data come from Andrea Frazzini's library: http://www.econ.yale.edu/~af227/data_library.htm; LIQ from Lubos Pastor: http://faculty.chicagobooth.edu/lubos.pastor/research/liq_data_1962_2014.txt; LIBOR, TERM, CREDIT, and TED from FRED: https://research.stlouisfed.org/fred2/ and VIX from the CBOE: http://www.cboe.com/micro/vix/historical.aspx. Finally, we thank Zhuo Chen, Ian Dew-Becker, and Grant Thomas Clayton for sharing FLS, VAR-SWAP, and LJV, respectively, and Lu Zhang for supplying the I/A and ROE data.

[29]The momentum strategy returns have a Fama and French (1993) three-factor alpha of 1.25%/month ($t = 2.22$) which is significantly positive and different from zero at 95% significance level during the period 1996:01-2013:12. Interestingly, the Hou et al. (2015) four-factor alpha is only 0.04%/month – which is not statistically significantly different from zero during this sample period. The four-factor model by Hou et al. (2015) seems to capture momentum strategy return risk profile well during this sample period.

6 Using the HMM to Forecast Value and Momentum Strategy Comovement

Up until now, we focused on demonstrating that a two-state HMM reliably forecasts momentum's tail events. In this section we show that the same state variable also forecasts comovement between momentum and value strategy returns, and captures the time-variation in this comovement better than other conditioning variables proposed in the literature.

Many asset management firms now combine value and momentum strategies as part of an integrated quantitative investment strategy.[30] At least in part, the motivation for doing so is the empirically observed negative correlation between value and momentum strategy returns; value serves as a natural hedge for momentum, and vice-versa. Asness et al. (2013) show that the performance of a combined value and momentum strategy is far stronger than the performance of the separate strategies as a result of the negative correlation between value and momentum returns.

This observed negative comovement is intriguing in that the unconditional loadings of value and momentum on the market are both close to zero. Thus, the negative correlation suggests that value and momentum both are exposed to a latent factor other than the market, but with opposite signs. Consistent with this, Asness et al. (2013) argue that this latent factor is associated with funding liquidity.

Our contribution to this literature is first to show that this negative correlation is present only in turbulent states. In addition we investigate the source of this negative comovement, and show that about half of the comovement is attributable to conditional loading on the market factor. In turbulent states, the market beta of the momentum strategy becomes negative and strong, while the beta of a value strategy is strongly positive, leading to a negative correlation.

---

[30]For example, Numeric Investors L.P. in Perold and Tierney (1997) utilize both value and an earnings momentum (Ball and Brown, 1968) metric in constructing their overall strategy.

However, even after orthogonalizing (conditionally) the strategy returns to the market, the strategies remain conditionally negatively correlated, and only in the turbulent state. We investigate the source of this residual conditional correlation, and find that it is related to liquidity variables, consistent with Asness et al. (2013).

6.1   Value-Momentum Strategy Return Correlations

For our analysis in this Section we define a zero-investment value portfolio VMG (Value-Minus-Growth) which is consistent with our momentum portfolio. VMG invests \$1 in the value-weighted portfolio of firms in VAL—the top-decile of BE/ME—and which shorts \$1 worth GRO—the corresponding bottom-decile portfolio—where NYSE decile breakpoints are used. Monthly returns from the VAL and GRO decile portfolios from Ken French's data library.[31]

A striking finding is that the well documented negative correlation between value and momentum strategy returns is present only when the market state is turbulent; in calm states the estimated correlation is statistically indistinguishable from zero. In Table 10 we show the results of an analysis where we first group the months in our sample into three equal-sized tertiles (High, Med, Low) based on the predicted probability of the hidden state being turbulent, $\Pr\left(S_t = Turbulent | \mathcal{F}_{t-1}\right)$. Panel A of Table 10 reports the correlation matrix for MOM, VMG, VAL (the high book-to-market portfolio), and GRO (the low book to market portfolio), for the three sets of months. The right-most correlation matrix is unconditional. A comparison of the MOM-VMG correlation across the three tertiles shows that the overall negative correlation (-0.44 (p-value=0.00)) is primarily driven by the negative correlation in High tertile: the MOM-VMG correlation in High tertile is -0.57 (p-value=0.00). In contrast, it becomes statistically indistinguishable from zero: -0.07 (p-value=0.20) and

---

[31]The construction of value and momentum factors by Asness et al. (2013) is slightly different. Refer to equation (1) on page 938 of their paper for more details. We have confirmed that we achieve qualitatively similar findings using either the Fama and French (1993) HML value portfolio or the AMP value portfolio. Table 15 in the Online Appendix summarizes the conditional comovement of MOM with these value factors.

-0.07 (p-value=0.21) for Med and Low tertiles, respectively.

We now proceed to explore the underlying drivers of the strong negative conditional correlation in turbulent states. One possible explanation of this pattern is that it results from the time-varying market betas of the MOM and VMG portfolios as pointed out by Kothari and Shanken (1997) and Grundy and Martin (2001). Specifically, in turbulent/high volatility times, when the market has generally fallen, the market beta of the momentum portfolio, which is short past losers, becomes strongly negative. Similarly, the value portfolio, which is long past losers, will have a strongly positive market beta in these states.

However, we do not find the past market return to be the only driver. To see this, we form three tertiles of Top, Middle (Mid), and Bottom (Btm) based on formation period return of the market using all sample months as well as within each state-probability tertile. The correlations between MOM and VMG returns from double sorts are reported in Panel A-1 and A-2 of Table 10.

First, we note that sorting on the predicted probability for the turbulent state shows a clearer separation on the correlation between MOM and VMG than sorting on the formation period market return. While the correlations in High, Med, and Low groups—formed by using the predicted probability for the turbulent state as a sorting variable—are -0.57, -0.07, and -0.07, respectively, those in Bottom, Middle, and Top groups – formed by using the formation period market return as a sorting variable – are -0.58, -0.44, and -0.13, respectively. We also find that the additional sorting on the formation period market returns does not change the broad pattern in the correlation between MOM and VMG conditional on the predicted probability for the turbulent state as shown in Panel A-1. Across subgroups of Bottom, Middle, and Top, the correlations between MOM and VMG are substantially more negative in the High tertile than in the Med or Low tertiles. That is, the concentration of negative correlation between MOM and VMG due to the high probability for the hidden state to be turbulent is a phenomenon different from, though related to, the effect of formation

period market returns. Further, we change the order of double sorts: we first sorted on the formation period market return, and then on the predicted probability of the underlying state being turbulent. Results are in Panel A-2. Note that even for the Middle tertile of the formation period market returns, further sorting on the probability of the underlying state being turbulent gives dispersion in the correlation between MOM and VMG. When the probability of the underlying state being turbulent is low, the correlation is not different from zero (0.03). In contrast, when the probability of the underlying state being turbulent is high, the correlation is strongly negative (-0.60).

A striking illustration of the effectiveness of the state probability in forecasting correlation comes from comparing Panel A-1 and A-2. In Panel A-1, when the predicted probability is in Medium tertile, the difference in the conditional correlation when moving from Bottom to Top tertile of the formation period return is 0.11. In contrast, in Panel A-2, when the formation period market return is in Middle tertile, moving from High to Low state probability, the conditional correlation changes by 0.63, i.e. higher by a factor of 6.

We next perform a more direct test of whether the negative conditional correlation between VMG and MOM in the turbulent state results from the common loading on the market. For all the months in each "state" (High, Med, and Low), we regress the MOM, VMG, VAL and GRO returns on the excess market return in univariate time-series regressions:

$$R_{p,t}^e = \alpha + \beta^0 R_{\text{MKT,t}}^e + e_p, \tag{17}$$

where $R_{\text{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with $p = \text{MOM}, \text{VMG}, \text{VAL}, \text{GRO}$. The slope coefficients, t-statistics and regression R^2s are reported in the top part of Panel B of Table 10. In the bottom part of Panel B we report the correlation of the sets of residuals from these four regressions. Consistent with our conjecture, when the probability that the hidden state is high, $\hat{\beta}_{\text{MOM}} = -0.89$ ($t = -6.83$), while $\hat{\beta}_{\text{VMG}} = 0.58$ ($t = 4.56$). Also, the VMG-MOM residual correlation of -0.25 is about

33

half as large as the correlation of -0.57 reported in Panel A.

Panel C of Table 10 performs essentially the same analysis as in Panel B, except that now we use the Henriksson and Merton (1981) specification in each regression:

$$R_{p,t}^e = \alpha + \beta^0 R_{\mathrm{MKT,t}}^e + \beta^+ \max\left(R_{\mathrm{MKT,t}}^e, 0\right) + e_p, \tag{18}$$

where $R_{\mathrm{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with $p = \mathrm{MOM, VMG, VAL, GRO}$. The same patterns in the residual correlations that we see in Panel B are also present here.

We see a correlation drop of about 50% (i.e., from -0.57 to only -0.25/-0.31 in Panels B and C, respectively) after controlling for market exposure, suggesting that about half of the VMG/MOM comovement is unexplained by the market. However, this statistic could be misleading. Since the correlation is equal to the covariance (comovement) divided by the product of the volatilities, it could be, for example, that approximately 90% of the covariance is explained by the market exposure, but the residual of each series variance falls by approximately 80% after controlling for the market.

To address this concern we can ask what fraction of the negative covariance between MOM and VMG is explained by market exposure. We do this in Table 11, where we report the residual covariance matrix. The comparison between Panel A and B of Table 11 reveals that 44% (=1-(35.58/64.05)) of negative covariance between MOM and VMG in the high group is explained by the opposite exposure to the market risk. Analyzing the negative comovement between MOM and VMG using the Henriksson-Merton specification makes the picture clearer. As can be seen from Panel C of Table 10, the difference between MOM and VMG on the exposure to the systematic risks in the High state is mostly due to the difference on the option-like features. For MOM, β^+ is -1.10 ($t = -3.78$). However, VMG has a significantly large β^+ of 0.94 ($t = 2.93$) but β^0 of VMG is very weak. Interestingly, the non-linearity of VMG in the High state is driven by the value side, not the growth side.

VAL has β^+ of is 0.85 (t=3.18) but the β^+ of GRO is only -0.09 (t=−1.36). Comparing the covariances between MOM and VMG in Panel A and C of Table 11, we can see that the linear and non-linear exposures to the market in MOM and VMG explain 55% (=1-(28.83/64.05)) of the negative covariance between MOM and VMG in the High tertile.

The residual negative comovement between MOM and VMG may be due to omitted factors such as liquidity (Asness et al., 2013). To examine this possibility, we regress the residuals from the CAPM and Henriksson-Merton (HM) specifications on proxies for economy-wide liquidity shocks. We consider (i) the funding liquidity shock measured by the innovations in AR(2) model of TED spread, the 3-month interbank LIBOR minus the 3-month T-bill rate, as well as (ii) the innovations in the aggregate market liquidity estimated by Pastor and Stambaugh (2003). We take the negative of the TED spread so that it represents the level of liquidity. Panel A (B) of Table 12 reports regression results using CAPM (HM) residuals. Consistent with the hypothesis of Asness et al. (2013), we find the positive (negative) dependence of the residual MOM (VMG) returns on liquidity shock proxies over the whole sample period, as shown in the last column of Table 12.

Interestingly, in the state-conditional regressions, the coefficient on the liquidity shocks are significant only in the High state. Furthermore, although the estimated coefficients are not always significant, the size of coefficients is much larger in the High state (first column) than the whole sample (last column). These findings highlight the importance of liquidity shocks that do not appear in MOM and VMG during times when the market is calm (2/3rd of the sample) and emerge when the market becomes turbulent. Economy-wide liquidity shocks affect stocks in the past loser and value portfolios in a similar manner. Since the MOM strategy is short past losers and VMG is long value stocks, liquidity shocks that emerge during turbulent market conditions drive a significant part of the negative correlation between value and momentum strategy returns.

6.2 Conditionally Hedging Value and Momentum Strategies

Asness et al. (2013) observe that a combined value/momentum portfolio, with constant weights on value and momentum, earns a higher unconditional Sharpe ratio than either portfolio alone. This improvement follows from the unconditional negative correlation between value and momentum returns. However, as we showed earlier, the correlation between value and momentum is almost zero except when the state proability is high. This suggests that a modification of the Asness et al. (2013) strategy by hedging MOM with VMG only in the high probability state will improve the Sharpe ratio further.

Table 13 gives the summary statistics of monthly returns for the following five portfolios over our sample period of 1044 months from 1927:01 to 2013:12: i) MOM, ii) VMG, iii) 50/50-static: a portfolio with constant weights of 0.5 on the MOM and VMG portfolios, iv) 50/50-dyn: a portfolio which puts the weights of 0.5 on both MOM and VMG in the High tertile, and the weights of 1.0 on MOM and 0.0 on VMG otherwise (Med and Low tertiles), v) Conditional MVE: a mean-variance efficient portfolio which puts weights of $w = (\gamma_s \mathbf{\Sigma}_s)^{-1} \boldsymbol{\mu}_s$ on VMG and MOM in each of two groups (High tertile / Med and Low tertiles) — that is, the portfolio weight in each group is conditional on *ex-post* mean and variance.

As documented in Asness et al. (2013), we find that the performance of MOM is greatly improved by adding VMG unconditionally: the SR of "50/50-static" is 0.76, which is much higher than those of MOM (0.52) and VMG (0.28). Also, the "50/50-static" portfolio experiences far fewer large losses than either MOM or VMG. For example, while there are 32 (9) large losses worse than -15% in MOM (VMG) portfolio from 1927:01 to 2013:12, "50/50-static" portfolio suffers from losses exceeding -15% only three times.

If the benefit of combining MOM with VMG is evenly spread out over our sample period, a portfolio of combining MOM with VMG only during sub sample periods would perform worse than the "50/50-static". However, we find the opposite to be true. By combining MOM with VMG only in High tertile, the SR improves substantially from 0.76 (50/50-

static) to 0.91 (50/50-dyn). The occurrence of monthly losses exceeding 10% decreases from 13 (50/50-static) to 8 (50/50-dyn). Interestingly, the performance of "50/50-dyn" portfolio is not much different from that of the "Conditional-MVE" portfolio. This implies that combining MOM with VMG only when the state probability is high is close to the optimal strategy.

7 Conclusion

There is a vast literature documenting that the simple strategy of buying past winners and selling past losers, commonly referred to as a momentum strategy, generates high average risk adjusted returns. However, such a strategy also experiences infrequent but large losses. We provide an explanation for the phenomenon, *i.e.*, why we see such large losses occurring at periodic but infrequent intervals. We show that the way momentum portfolios are formed embeds features that resemble a written call option on the market portfolio into the momentum strategy returns. These features become accentuated in prolonged bear markets, when the market is volatile due to increased financial and operating leverage. These dynamics lead to large momentum strategy losses when the market recovers.

The intuition for the optionality follows Merton (1974): after large negative market returns the effective leverage of the firms on the short side of the momentum strategy (the past-loser firms) becomes extreme. As the firm values fall, driven both by the market and other factors, the common shares move from being deep in-the-money call options on the firm's underlying assets to at- or even out-of-the-money options, and thus start to exhibit the convex payoff structure associated with call options: the equity value changes little in response to even large down moves in the underlying firm value, but moves up dramatically in response to up moves. Thus, when the values of the firms in the loser portfolio increase—proxied by positive returns on the market portfolio—the convexity in the option payoff results

in outsized gains in the past loser portfolio. Since the momentum portfolio is short these loser firms, this results in dramatic losses for the overall long-short momentum portfolio.

High leverage of the past-loser portfolio is the driver of the tail-risk of the momentum strategies in our model. Consistent with this we show that, just preceding the five worst momentum crash months in the 1964-2013 period, the average financial leverage of the past loser portfolio was 47.2, compared with an unconditional average over this period of 3.97. However *effective leverage*—both financial and operating leverage effect—drives the optionality of the past loser portfolio, and effective leverage is difficult to measure. Consequently we proceed to estimate the time-varying tail risk of momentum strategies using the two-state hidden Markov model (HMM) where the embedded option-like features of momentum strategy returns become accentuated in the hidden *turbulent* state. Empirically, we find that the behavior of both the market and the momentum portfolio are consistent with this model. When the economy is in the latent *turbulent* state, the levels of market and momentum strategy volatility are more than double their values in the calm state. In the *tubulent* state, the option-like features of the momentum strategy are pronounced; in the *calm* state they are dramatically attenuated.

We find that momentum crashes tend to occur more frequently during months in which the hidden state is more likely to be turbulent. The turbulent state occurs infrequently in the sample: the probability that the hidden state is turbulent exceeds 80% in only 155 of the 1044 months in our 1927:01-2013:12 sample. Yet in each of the 13 severe loss months, the *ex-ante* probability that the hidden state is turbulent exceeds 80 percent. Interestingly, the average momentum strategy return during those 155 months is only -1.01% per month.

We derive the conditional alpha of the momentum strategy for a given month based on the information available untill the end of the previous month using the HMM return generating process for the momentum strategy returns and market excess returns and the price of call options on the market and the risk free rate. Over the 216 month period for which we have

call option prices (1996:01-2013:12) the average conditional alpha is 1.17%/month, which is significantly positive. However, the conditional alpha is negative during 49 out of the 216 months and significantly negative for three months 2008:11-2008:12 and 2009:03 of the financial crisis period.

We show that QML estimator of HMM parameters need not to be consistent when the wrong likelihood is maximized. We find that the normally distributed residuals for momentum strategy returns and Student's t-distributed (d.f.=5) residuals for market excess returns best describe the data. Our HMM model has the fewest number of false positives in predicting momentum crashes when compared to models using historical realized volatility, GARCH or past market returns.

We show that the same turbulent state probability that forecasts momentum crashes also reliably forecasts the value-momentum strategy return correlation. Interestingly, when the *ex-ante* turbulent state probability is low, the conditional value-momentum strategy return correlation is statistically indistinguishable from zero. However, when the turbulent state probability is at the highest levels, the conditional correlation is below -0.5.

We find that value is a good hedge for momentum only during turbulent times. During those time periods, half the negative covariance between value and momentum arises from nonlinear exposure to the market factor. The other half of the negative covariance is due to exposure to an economy wide pervasive risk factor factor that emerges during turbulent times. There is some empirical support for the view that this pervasive risk factor is related to liquidity risk identified by Asness et al. (2013).

References

Asness, Clifford S., Andrea Frazzini, and Lasse H. Pedersen, 2014, Quality minus junk, Working paper, SSRN.

Asness, Clifford S., Tobias J. Moskowitz, and Lasse Heje Pedersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Badrinath, Swaminathan G, and Sunil Wahal, 2002, Momentum trading by institutions, *Journal of Finance* 57, 2449–2478.

Ball, Ray, and Philip Brown, 1968, An empirical examination of accounting income numbers, *Journal of Accounting Research* 159–178.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barroso, Pedro, and Pedro Santa-Clara, 2015, Momentum has its moments, *Journal of Financial Economics* 116, 111–120.

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, Viktor Todorov, and Lai Xu, 2015, Tail risk premia and return predictability, *Journal of Financial Economics* 118, 113–134.

Breen, William, Lawrence R Glosten, and Ravi Jagannathan, 1989, Economic significance of predictable variations in stock index returns, *Journal of Finance* 44, 1177–89.

Chabot, Benjamin, Eric Ghysels, and Ravi Jagannathan, 2014, Momentum trading, return chasing and predictable crashes, Working paper.

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

Chen, Zhuo, and Andrea Lu, 2015, A market-based funding liquidity measure, Working paper, Tsinghua University and University of Melbourne.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 1998, Investor psychology and security market under- and overreactions, *Journal of Finance* 53, 1839–1885.

Daniel, Kent, and Tobias J. Moskowitz, 2016, Momentum crashes, *Journal of Financial Economics* 122, 221 – 247.

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

Dew-Becker, Ian, Stefano Giglio, Anh Le, and Marius Rodriguez, 2015, The price of variance risk, Working paper, Northwestern University, University of Chicago, Penn State University, and the Federal Reserve Board.

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., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Frazzini, Andrea, and Lasse Heje Pedersen, 2014, Betting against beta, *Journal of Financial Economics* 111, 1–25.

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

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

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

Grinblatt, Mark, and Sheridan Titman, 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.

Han, Bing, Dong Hong, and Mitch Warachka, 2009, Forecast accuracy uncertainty and momentum, *Management Science* 55, 1035–1046.

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.

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

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: an investment approach, *Review of Financial Studies* 28, 650–705.

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

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

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

Koijen, Ralph, Juan Carlos Rodriguez, and Alessandro Sbuelz, 2009, Momentum and mean reversion in strategic asset allocation, *Management Science* 55, 1199–1213.

Kothari, S.P., and Jay Shanken, 1997, Book-to-market, dividend yield, and expected market returns: A time-series analysis, *Journal of Financial Economics* 44, 169–203.

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

Lettau, Martin, Matteo Maggiori, and Michael Weber, 2014, Conditional risk premia in currency markets and other asset classes, *Journal of Financial Economics* 114, 197 – 225.

Liu, Laura Xiaolei, and Lu Zhang, 2008, Momentum profits, factor pricing, and macroeconomic risk, *Review of Financial Studies* 21, 2417–2448.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of risk and return in risk arbitrage, *Journal of Finance* 56, 2135–2175.

Moreira, Alan, and Tyler Muir, 2015, Volatility managed portfolios, Yale SOM Working Paper.

Pastor, Lubos, and Robert F. Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Perold, Andre F, and Brian J Tierney, 1997, Numeric investors l.p., HBS No. 9-298-012, Harvard Business School Publications, Boston.

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

Swaminathan, Bhaskaran, 2010, Quantitative money management: A practical application of behavioral finance, Working paper.

White, Halbert, 1980, A heteroskedasticity-consistent covariance estimator and direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Table 1: SUMMARY STATISTICS OF MOMENTUM STRATEGY RETURNS

Panel A reports the mean, standard deviation (SD), annualized Sharpe ratio (SR), skewness (skew), kurtosis (kurt), maximum (max), and minimum (min) of momentum strategy returns (MOM) along with those of market excess returns (Mkt-Rf), and scaled market excess returns (Mkt-Rf*) with the standard deviation equal to that of momentum strategy returns. Panel B reports the average risk adjusted monthly return (alpha), calculated as the intercept from time series regressions of the MOM return on the Market along with the corresponding risk exposures (betas). The sample period is 1927:01-2013:12. The t-statistics are computed using the heteroskedasticity consistent covariance estimator (White, 1980). The mean, SD, max and min in Panel A and α in Panel B are reported in percentage per month.

PANEL A: SUMMARY STATISTICS							
	mean	SD	SR	skew	kurt	max	min
MOM	1.18	7.94	0.52	-2.43	21.22	26.18	-79.57
Mkt-Rf	0.64	5.43	0.41	0.16	10.35	38.04	-29.10
Mkt-Rf*	0.94	7.94	0.41	0.16	10.35	55.74	-42.64

PANEL B: RISK ADJUSTED MOM RETURNS			
	α	$\beta_{\text{Mkt−Rf}}$	$Adj.R^2$
ESTIMATE	1.52	-0.52	0.13
(t-stat)	(7.10)	(-4.82)	

Wooldridge, Jeffrey, 1986, Estimation and inference for dependent processes, in R. F. Engle, and D. McFadden, eds., *Handbook of Econometrics*, volume 4, first edition, chapter 45, 2639–2738 (Elsevier).

Table 2: Market Conditions during Momentum Crashes

Panel A presents the momentum strategy returns (R_{MOM}), and the excess returns of winner portfolio, loser portfolio and market portfolio, denoted by R_{WIN}^e, R_{LOS}^e and Mkt-Rf, respectively, during months with the momentum crashes worse than -20% during 1927:01-2013:12 along with the breakpoints for the winner and loser portfolios, i.e., the threshold values for the cumulative returns over the measurement period from month $t-12$ to $t-2$, i.e., (*12-2 Ret*) for entering the winner and loser portfolios , and the ratio of the book value of debt to the market value of equity (BD/ME) of the winner and loser portfolios, the cumulative market returns in percentage during the 36 and 12 months preceding the month in which the momentum portfolios are formed, and the realized volatility of daily market returns during the 12 and 6 months preceding the month in which the momentum portfolios are formed. Sample averages of the variables across thirteen months in which the momentum crashes were realized are reported in Panel B and the averages of those variables across all available data are reported in Panel C. The book value of debt (BD) is available from 1964 onwards. Realized volatility is computed as the square root of the sum of squared daily returns and reported as annualized percentage. All of the variables except BD/ME are reported in percentage.

Date	R_{MOM}	R_{WIN}^e	R_{LOS}^e	Mkt-Rf	Winner Portfolio		Loser Portfolio		Past Mkt Ret		Past Mkt RV	
					Break-points	BD/ME	Break-points	BD/ME	36 Mos.	12 Mos.	12 Mos.	6 Mos.
Panel A: Momentum Crash Months												
1931:06	−29.03	8.17	37.20	13.79	−0.64	n.a.	−74.07	n.a.	−36.67	−45.68	22.74	20.90
1932:07	−60.37	13.95	74.32	33.60	−33.12	n.a.	−88.35	n.a.	−81.52	−65.87	41.73	40.23
1932:08	−79.57	14.36	93.93	36.46	−30.61	n.a.	−86.25	n.a.	−76.45	−51.19	41.60	38.91
1932:11	−22.68	−20.83	1.85	−5.61	50.00	n.a.	−50.00	n.a.	−67.25	−27.00	44.65	50.32
1933:04	−41.94	28.77	70.71	38.04	55.17	n.a.	−54.55	n.a.	−72.52	−12.66	45.55	39.33
1933:05	−28.03	19.27	47.30	21.38	114.00	n.a.	−41.94	n.a.	−61.25	46.97	45.32	40.36
1938:06	−33.34	10.45	43.79	23.72	−9.16	n.a.	−68.93	n.a.	8.60	−39.09	32.45	29.26
1939:09	−44.57	7.92	52.49	16.96	51.22	n.a.	−33.33	n.a.	−16.29	−0.96	19.64	19.54
2001:01	−41.97	−6.94	35.03	3.12	94.09	0.08	−55.17	0.68	37.67	−11.58	24.48	22.15
2002:11	−20.40	2.12	22.52	5.96	64.42	0.12	−48.34	2.32	−30.71	−13.63	23.99	30.04
2009:03	−39.31	4.81	44.12	8.95	7.25	0.06	−79.52	70.55	−38.37	−42.63	42.11	56.03
2009:04	−45.89	−0.13	45.76	10.19	−1.76	0.07	−82.44	106.89	−34.05	−37.00	43.60	56.15
2009:08	−24.80	0.21	25.01	3.32	15.98	0.10	−66.10	55.57	−15.22	−18.90	45.14	32.77
Panel B: Averages Across Momentum Crash Months												
	−39.38	6.32	45.69	16.14	28.99	0.09	−63.77	47.20	−37.23	−24.56	36.39	36.62
Panel C: Average Across All Available Sample Months												
	1.18	1.24	0.06	0.64	76.08	0.10	−23.24	3.97	38.37	11.95	14.82	14.58

Table 3: Option-like Feature of Momentum Returns and Market Conditions

We partition the months in our sample into three groups on the basis of the cumulative market return during the 36 months immediately preceding the momentum portfolio formation date. The 'High' ('Low') group consists of all months in which this variable is in the top (bottom) 20th percentile. The rest of the months are classified as 'Medium'. We estimate equation (1): using ordinary least squares for the months within each group, and report the results in Panel A. The dependent variable is either: the momentum strategy returns (R_{MOM}), or the returns of the winner or loser portfolio in excess of risk free return (R_{WIN}^e and R_{LOS}^e). For comparison, in Panel B we report the estimates for the CAPM, without the exposure to the call option on the market in (1). Panel C counts the number of momentum losses worse than 20% within each group. Panel D reports the skewness of $R_{p,t}^e$ with that of estimated ε of (1). α is reported in percentage per month. The t-statistics are computed using the heteroskedasticity-consistent covariance estimator by White (1980). The sample period is 1929:07-2013:12.

	State Variable: Past 36 Months Market Returns								
	Low			Medium			High		
R_p^e:	R_{MOM}	R_{WIN}^e	R_{LOS}^e	R_{MOM}	R_{WIN}^e	R_{LOS}^e	R_{MOM}	R_{WIN}^e	R_{LOS}^e
A: Henriksson-Merton Estimates									
α	3.00	1.12	-1.88	2.62	1.03	-1.59	0.59	0.65	0.06
$t(\alpha)$	(3.28)	(2.88)	(-2.93)	(5.20)	(3.44)	(-6.03)	(1.04)	(1.82)	(0.17)
β^0	-0.44	0.97	1.41	-0.08	1.27	1.35	0.20	1.40	1.19
$t(\beta^0)$	(-3.14)	(12.45)	(16.60)	(-0.42)	(11.80)	(13.23)	(1.89)	(17.93)	(13.45)
β^+	-1.01	-0.29	0.72	-0.49	-0.26	0.23	0.26	-0.12	-0.38
$t(\beta^+)$	(-3.17)	(-2.02)	(3.60)	(-1.51)	(-1.37)	(1.39)	(0.94)	(-0.78)	(-1.92)
$Adj.R^2$	0.48	0.77	0.83	0.06	0.71	0.72	0.06	0.81	0.62
B: CAPM Estimates									
α	0.05	0.27	0.22	1.81	0.60	-1.21	1.08	0.41	-0.67
$t(\alpha)$	(0.08)	(1.05)	(0.45)	(7.42)	(4.62)	(-7.42)	(2.74)	(1.90)	(-2.45)
β	-1.02	0.80	1.82	-0.32	1.14	1.47	0.31	1.35	1.04
$t(\beta)$	(-6.60)	(13.38)	(18.01)	(-3.40)	(21.17)	(27.93)	(3.45)	(30.06)	(15.37)
$Adj.R^2$	0.43	0.76	0.82	0.05	0.71	0.71	0.06	0.81	0.62
C: Number of Momentum Losses worse than -20%									
	11			2			0		
D: Conditional Skewness									
R_p^e	-2.33	-0.21	1.74	-0.98	-0.59	0.07	0.17	-0.73	-0.61
ε_p	-0.48	-0.59	0.59	-0.76	-0.21	0.94	-0.12	1.07	0.81

Table 4: OPTION-LIKE FEATURE OF MOMENTUM RETURNS DURING DOT-COM CRASH

We estimate equation (1) with the momentum strategy return (R_{MOM}) and the winner and loser portfolio excess returns (R^e_{WIN} and R^e_{LOS}) as a candidate dependent variable. We use 36 monthly data on returns during 2000:01-2002:12. α is reported in percentage per month. The t-statistics are computed using the heteroskedasticity-consistent covariance estimator by White (1980).

$R^e_p:$	R_{MOM}	R^e_{WIN}	R^e_{LOS}
α	3.41	1.57	-1.84
$t(\alpha)$	(0.92)	(0.85)	(-0.71)
β^0	-0.42	1.25	1.67
$t(\beta^0)$	(-0.71)	(3.06)	(4.31)
β^+	-1.35	-0.54	0.82
$t(\beta^+)$	(-1.26)	(-0.81)	(1.11)

Table 5: MAXIMUM LIKELIHOOD ESTIMATES OF HMM PARAMETERS

We maximize the likelihood of data with the assumption that $\varepsilon_{\mathrm{MOM},t}$ in (2) is drawn from a standard normal distribution and $\varepsilon_{\mathrm{MKT},t}$ in (3) is drawn from a Student's t-distribution with d.f.=5. The parameters are estimated using data for the period 1927:01-2013:12. α, σ_{MOM}, and σ_{MKT} are reported in percentage per month.

	HIDDEN STATE			
	$S_t = \mathrm{Calm}(C)$		$S_t = \mathrm{Turbulent}(T)$	
PARAMETER	ESTIMATES	(T-STAT)	ESTIMATES	(T-STAT)
α (%)	1.95	(7.35)	4.05	(3.82)
β^0	0.34	(3.79)	-0.32	(-1.90)
β^+	-0.46	(-3.27)	-1.14	(-4.63)
σ_{MOM} (%)	4.31	(17.21)	11.02	(20.56)
μ	1.11	(8.13)	-0.38	(-0.74)
σ_{MKT} (%)	4.04	(26.74)	8.36	(15.80)
$\Pr\left(S_t = s_{t-1} \mid S_{t-1} = s_{t-1}\right)$	0.98	(9.52)	0.94	(10.34)

Table 6: Momentum and Market Excess Returns: Sample Moments vs HMM-implied Moments

We compare the HMM-implied moments of momentum strategy returns and market excess returns with the corresponding moments in our sample. After we estimate HMM parameters, we generate $\varepsilon_{\text{MOM},t}$ and $\varepsilon_{\text{MKT},t}$ in our HMM specification of (2)) and (3)) using monte carlo simulation from various combinations of normal and Student's t-distributions. Then, we construct a 1044 month-long time series of momentum strategy and market excess returns using HMM specification and compute their first four moments. We then repeat this exercise 10,000 times to obtain the distribution of the first four momentums.

	Momentum Strategy Returns: $R_{\text{MOM},t}$						Market Excess Returns: $R^e_{\text{MKT},t}$					
	Realized Moments	Quantiles (%) of Simulated Moments					Realized Moments	Quantiles (%) of Simulated Moments				
		0.5	2.5	50	97.5	99.5		0.5	2.5	50	97.5	99.5
Panel A: Normal ($\varepsilon_{\text{MOM},t}$) and Student's t ($\varepsilon_{\text{MKT},t}$)												
mean	1.18	0.49	0.66	1.21	1.69	1.81	0.64	0.04	0.23	0.67	1.08	1.19
std.dev	7.94	6.18	6.62	8.20	9.87	10.20	5.43	4.61	4.87	5.69	6.61	6.96
skewness	-2.43	-5.83	-2.45	-0.60	0.00	0.16	0.16	-3.51	-1.81	-0.27	1.58	4.37
kurtosis	21.22	5.28	5.63	7.94	29.79	79.53	10.35	5.12	5.54	8.45	40.00	80.70
Panel B: Normal ($\varepsilon_{\text{MOM},t}$) and Normal ($\varepsilon_{\text{MKT},t}$)												
mean	1.18	0.46	0.63	1.13	1.61	1.72	0.64	0.17	0.28	0.65	0.99	1.09
std.dev	7.94	6.26	6.59	7.76	8.95	9.34	5.43	4.58	4.76	5.39	6.05	6.27
skewness	-2.43	-1.40	-1.20	-0.58	0.01	0.24	0.16	-0.88	-0.74	-0.34	0.08	0.22
kurtosis	21.22	5.83	6.27	8.09	11.51	13.05	10.35	4.52	4.76	5.79	7.47	8.37
Panel C: Student's t ($\varepsilon_{\text{MOM},t}$) and Normal ($\varepsilon_{\text{MKT},t}$)												
mean	1.18	0.62	0.79	1.28	1.75	1.92	0.64	0.10	0.25	0.66	1.02	1.11
std.dev	7.94	5.88	6.33	7.78	9.35	9.96	5.43	4.30	4.56	5.37	6.23	6.46
skewness	-2.43	-3.27	-1.78	-0.50	0.80	2.19	0.16	-0.94	-0.80	-0.39	0.03	0.16
kurtosis	21.22	6.17	6.80	9.86	28.46	64.22	10.35	4.40	4.68	5.76	7.51	8.32
Panel D: Student's t ($\varepsilon_{\text{MOM},t}$) and Student's t ($\varepsilon_{\text{MKT},t}$)												
mean	1.18	0.78	0.93	1.42	1.89	2.05	0.64	0.15	0.30	0.75	1.13	1.24
std.dev	7.94	5.56	6.03	7.67	9.53	10.17	5.43	4.33	4.58	5.42	6.45	6.87
skewness	-2.43	-4.36	-2.25	-0.49	0.93	2.38	0.16	-3.34	-1.72	-0.29	1.15	2.74
kurtosis	21.22	6.27	6.87	10.66	36.30	82.36	10.35	5.02	5.54	8.49	32.09	75.30

Table 7: EXTREME LOSSES/GAINS CONDITIONAL ON $\Pr\left(S_t = \text{Turbulent}|\mathcal{F}_{t-1}\right)$

This table presents the fraction of the total number of extreme losses/gains greater than a given value that occur when $\Pr\left(S_t = \text{Turbulent}|\mathcal{F}_{t-1}\right)$ is larger than a given threshold. The sample period is 1927:01-2013:12.

$\Pr\left(S_t = T|\mathcal{F}_{t-1}\right)$	# EXTREME LOSSES DURING TURBULENT MONTHS /# EXTREME LOSSES IN THE SAMPLE					# of
IS MORE THAN	$\leq -20\%$	$\leq -17.5\%$	$\leq -15\%$	$\leq -12.5\%$	$\leq -10\%$	MONTHS
90%	10/13	15/21	18/32	22/37	25/56	108
80%	13/13	19/21	23/32	27/37	30/56	155
70%	13/13	19/21	23/32	27/37	31/56	178
60%	13/13	19/21	23/32	27/37	33/56	206
50%	13/13	20/21	25/32	29/37	37/56	236
40%	13/13	21/21	27/32	31/37	40/56	265
30%	13/13	21/21	28/32	32/37	42/56	285
20%	13/13	21/21	29/32	33/37	43/56	317
10%	13/13	21/21	31/32	35/37	47/56	385

$\Pr\left(S_t = T|\mathcal{F}_{t-1}\right)$	# EXTREME GAINS DURING TURBULENT MONTHS /# EXTREME GAINS IN THE SAMPLE					# of
IS MORE THAN	$\geq 20\%$	$\geq 17.5\%$	$\geq 15\%$	$\geq 12.5\%$	$\geq 10\%$	MONTHS
90%	5/12	6/15	10/28	17/45	26/74	108
80%	8/12	10/15	16/28	24/45	33/74	155
70%	8/12	10/15	16/28	27/45	37/74	178
60%	8/12	10/15	16/28	27/45	38/74	206
50%	9/12	11/15	19/28	30/45	43/74	236
40%	9/12	11/15	20/28	31/45	44/74	265
30%	9/12	11/15	20/28	32/45	49/74	285
20%	10/12	12/15	21/28	33/45	52/74	317
10%	11/12	14/15	23/28	36/45	60/74	385

Table 8: FALSE POSITIVES IN PREDICTING MOMENTUM CRASHES

We compare the number of false positives in predicting momentum crashes across different tail risk measures. The number of false positives of a given tail risk measure is computed as follows. Suppose we classify months in which momentum strategy returns lost more than a threshold X. Let Y denote the lowest value attained by a given tail risk measure during those momentum crash months. During months when the tail risk measure is above the threshold level of Y, we count the number of months when momentum crashes did not occur and we denote it as the number of false positives. We consider X=10%, 20%, 30%, 40%. In Panel A, we use $\Pr(S_t = \text{Turbulent}|\mathcal{F}_{t-1})$ as a tail risk measure. The results in Panel A-1 are from our original HMM model specified in (2), (3) and (4). To emphasize the importance of option-like feature $\beta^+(S_t)$ in (2), we impose the restriction $\beta^+(S_t) = 0$ and report the associated results in Panel A-2. In Panel B, we use various estimates of the volatility of momentum strategy returns as tail risk measures. Specifically, we estimate the volatility of the momentum strategy returns using GARCH (1,1) and realized volatility of daily momentum strategy returns over the previous 3, 6, 12, and 36 months. In Panel C, we use the volatility of the market return estimated using GARCH(1,1) and realized volatility of the daily market return during the preceding 3, 6, 12, and 36 months as tail risk measures. In Panel D, we use the market return during the preceding 3, 6, 12 and 36 month windows as tail risk measures.

	MOMENTUM CRASH THRESHOLD (-X)			
TAIL RISK MEASURE	$\leq -40\%$	$\leq -30\%$	$\leq -20\%$	$\leq -10\%$
PANEL A: HMM				
A-1: MAIN SPECIFICATION				
$\Pr(S_t = \text{T}|\mathcal{F}_{t-1})$	121	119	114	902
A-2: WITHOUT THE OPTION-LIKE FEATURE $\beta^+(S_t) = 0$				
$\Pr(S_t = \text{T}|\mathcal{F}_{t-1})$	157	155	150	891
PANEL B: MOMENTUM STRATEGY RETURNS VOLATILITY				
GARCH(1,1)	263	261	256	829
RV(3 MONTHS)	234	232	227	922
RV(6 MONTHS)	194	192	187	892
RV(12 MONTHS)	154	152	147	866
RV(36 MONTHS)	180	178	173	951

Table 8 – continued from previous page

PANEL C: MARKET RETURNS VOLATILITY				
GARCH(1,1)	188	186	181	809
RV(3 MONTHS)	166	164	159	889
RV(6 MONTHS)	183	181	176	920
RV(12 MONTHS)	191	189	184	796
RV(36 MONTHS)	179	177	172	858
PANEL D: PAST MARKET RETURNS				
3 MONTHS	618	616	980	948
6 MONTHS	131	129	918	944
12 MONTHS	249	247	966	938
36 MONTHS	491	489	484	930

Table 9: SYSTEMATIC RISK IN MOMENTUM STRATEGY RETURNS

This table presents the results of regressing $\epsilon_{\text{MOM},t}$ in (15) on various systematic risk factors, $\epsilon_{\text{MOM},t} = \text{intercept} + \text{coeff} \times \text{systematic factor}_t + e_t$. Results are obtained by using the data from 1996:01 to 2013:12 (216 months) where we can reconstruct $\epsilon_{\text{MOM},t}$ from the market prices of call option on S&P 500 from OptionMetrics. Details on systematic factors are described in the main text.

PANEL A: GENERAL FACTORS						
systematic factor	coeff	t(coeff)	$R^2(\%)$	First Month	Last Month	N
MKT	-0.12	-0.78	0.41	1996:01	2013:12	216
SMB	0.25	0.94	1.08	1996:01	2013:12	216
HML	-0.52	-1.97	4.36	1996:01	2013:12	216
RMW	-0.03	-0.09	0.01	1996:01	2013:12	216
CMA	-0.04	-0.09	0.01	1996:01	2013:12	216
I/A	-0.27	-0.59	0.48	1996:01	2013:12	216
ROE	0.99	2.99	13.00	1996:01	2013:12	216
QMJ	0.51	1.68	3.66	1996:01	2013:12	216
PANEL B: LIQUIDITY RELATED FACTORS						
systematic factor	coeff	t(coeff)	$R^2(\%)$	First Month	Last Month	N
LIQ	0.26	1.32	1.59	1996:01	2013:12	216
FLS	0.07	0.56	0.71	1996:01	2012:10	202
BAB	0.21	0.82	1.14	1996:01	2012:03	195
Δ LIBOR	0.05	1.23	1.73	1996:01	2013:12	216
Δ TERM	-0.04	-1.39	1.18	1996:01	2013:12	216
Δ CREDIT	0.03	0.86	0.22	1996:01	2013:12	216
Δ TED	-0.02	-0.49	0.13	1996:01	2013:12	216
PANEL C: TAIL RISK RELATED FACTORS						
systematic factor	coeff	t(coeff)	$R^2(\%)$	First Month	Last Month	N
VAR-SWAP 1M	0.00	-0.52	0.06	1996:02	2013:09	212
VAR-SWAP 3M	0.01	0.70	0.09	1996:02	2013:08	210
VAR-SWAP 6M	0.02	1.41	0.44	1996:08	2013:08	203
VAR-SWAP 12M	0.01	0.57	0.09	1997:03	2013:08	193
Δ VIX	0.00	1.03	0.39	1996:01	2013:12	216
Δ LJV	-0.77	-0.42	0.09	1996:01	2013:12	216

Table 10:   Realized Correlation, CAPM-implied Correlation, and HM-implied Correlation Conditional on Predicted Probability for the Turbulent State

This table represents the conditional nature of the correlation between MOM and VMG. We group the months in our sample into three equal-sized tertiles (High, Med, Low) based on the predicted probability of the hidden state being turbulent, $\Pr\left(S_t = Turbulent | \mathcal{F}_{t1}\right)$. The state probability breakpoints for the three tertiles are 14% and 3%. Panel A reports the correlation matrix for MOM, VMG, VAL (the high book-to-market portfolio), and GRO (the low book to market portfolio), for the three sets of months. The right-most correlation matrix is unconditional. To separate the effect of the past market returns on the correlation, we form three tertiles of Top, Middle (Mid), and Bottom (Btm) based on formation period return of the market using all sample months as well as within each state-probability tertile. When we use all 1044 months, the breakpoints for formation period market returns are 19.22% and 4.49% . The correlations from double sorts are reported in Panel A-1 and A-2. In Panel B, we estimate the following time series regressions: $R_{p,t}^e = \alpha + \beta^0 R_{\mathrm{MKT,t}}^e + e_p$, where $R_{\mathrm{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with $p = $ MOM, VMG, VAL, GRO. In Panel C, we use the the Henriksson and Merton (1981) specification in each regression: $R_{p,t}^e = \alpha + \beta^0 R_{\mathrm{MKT,t}}^e + \beta^+ \max\left(R_{\mathrm{MKT,t}}^e, 0\right) + e_p$, where $R_{\mathrm{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with $p = $ MOM, VMG, VAL, GRO. The t-statistics are computed using the heteroskedasticity consistent covariance estimator (White, 1980) and reported in parentheses in Panel B and C.

Panel A: Realized Correlation conditional on Predicted Probability for the Turbulent State

	High				Med				Low				All		
	MOM	VMG	VAL		MOM	VMG	VAL		MOM	VMG	VAL		MOM	VMG	VAL
VMG	-0.57			VMG	-0.07			VMG	-0.07			VMG	-0.44		
VAL	-0.66	0.84		VAL	0.04	0.68		VAL	0.00	0.63		VAL	-0.48	0.79	
GRO	-0.47	0.28	0.75	GRO	0.12	-0.03	0.72	GRO	0.07	-0.19	0.64	GRO	-0.28	0.14	0.72

A-1: Corr(MOM,VMG) by double sorts on 1) predicted probability and 2) formation period return of market

	High				Med				Low				All		
	Btm	Mid	Top		Btm	Mid	Top		Btm	Mid	Top		Btm	Mid	Top
	-0.69	-0.54	-0.30		-0.14	-0.05	-0.03		-0.10	-0.20	0.07		-0.58	-0.44	-0.13

A-2: Corr(MOM,VMG) by double sorts on 1) formation period return of market and 2) predicted probability

	Btm				Mid				Top				All		
	High	Med	Low		High	Med	Low		High	Med	Low		High	Med	Low
	-0.63	-0.53	-0.17		-0.60	-0.15	0.03		-0.22	-0.10	0.07		-0.57	-0.07	-0.07

Table 10 – continued from previous page

PANEL B: CAPM-IMPLIED RESIDUAL CORRELATION CONDITIONAL ON PREDICTED PROBABILITY FOR THE TURBULENT STATE

	High β	R^2		Med β	R^2		Low β	R^2		All β	R^2
MOM	-0.89	0.30	MOM	0.08	0.01	MOM	0.05	0.00	MOM	-0.52	0.13
	(-6.83)			(1.07)			(0.81)			(-4.82)	
VMG	0.58	0.21	VMG	0.23	0.05	VMG	0.12	0.01	VMG	0.43	0.13
	(4.56)			(3.45)			(1.72)			(4.91)	
VAL	1.58	0.74	VAL	1.28	0.72	VAL	1.15	0.66	VAL	1.44	0.71
	(14.29)			(23.72)			(21.36)			(18.87)	
GRO	0.99	0.92	GRO	1.05	0.89	GRO	1.03	0.85	GRO	1.01	0.90
	(43.19)			(43.89)			(40.51)			(63.85)	

	e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}
e_{VMG}	-0.25			e_{VMG}	0.02			e_{VMG}	0.01			e_{VMG}	-0.13		
e_{VAL}	-0.47	0.40		e_{VAL}	0.06	0.19		e_{VAL}	0.04	0.09		e_{VAL}	-0.30	0.31	
e_{GRO}	-0.53	0.45	0.83	e_{GRO}	0.07	0.21	0.80	e_{GRO}	0.04	0.10	0.75	e_{GRO}	-0.34	0.34	0.80

Table 10 – continued from previous page

PANEL C: HM-IMPLIED RESIDUAL CORRELATION CONDITIONAL ON PREDICTED PROBABILITY FOR THE TURBULENT STATE																	
	High				Med				Low				All			
	β^0	β^+	R^2		β^0	β^+	R^2		β^0	β^+	R^2		β^0	β^+	R^2
MOM	-0.31	-1.10	0.35	MOM	0.19	-0.21	0.01	MOM	0.26	-0.42	0.02	MOM	-0.06	-0.89	0.17
	(-2.28)	(-3.78)			(1.44)	(-0.92)			(2.74)	(-2.01)			(-0.52)	(-3.12)	
VMG	0.08	0.94	0.28	VMG	0.57	-0.66	0.09	VMG	0.06	0.12	0.01	VMG	0.17	0.50	0.15
	(0.63)	(2.93)			(5.66)	(-4.00)			(0.60)	(0.54)			(1.85)	(1.95)	
VAL	1.12	0.85	0.76	VAL	1.54	-0.51	0.73	VAL	1.04	0.21	0.66	VAL	1.19	0.48	0.72
	(10.20)	(3.18)			(20.04)	(-3.97)			(14.31)	(1.22)			(14.95)	(2.22)	
GRO	1.04	-0.09	0.93	GRO	0.97	0.14	0.89	GRO	0.98	0.09	0.85	GRO	1.02	-0.02	0.90
	(33.57)	(-1.36)			(26.86)	(2.15)			(24.39)	(1.23)			(45.34)	(-0.32)	
	e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}		e_{MOM}	e_{VMG}	e_{VAL}
e_{VMG}	-0.31			e_{VMG}	0.03			e_{VMG}	0.00			e_{VMG}	-0.16		
e_{VAL}	-0.51	0.44		e_{VAL}	0.07	0.21		e_{VAL}	0.03	0.09		e_{VAL}	-0.32	0.32	
e_{GRO}	-0.52	0.44	0.82	e_{GRO}	0.07	0.20	0.79	e_{GRO}	0.04	0.10	0.75	e_{GRO}	-0.34	0.34	0.80

Table 11: Covariance decomposition of MOM and VMG returns / residuals

This table represents the conditional nature of the covariance between MOM and VMG. We group the months in our sample into three equal-sized tertiles (High, Med, Low) based on the predicted probability of the hidden state being turbulent, $\Pr\left(S_t = Turbulent | \mathcal{F}_{t1}\right)$. Panel A reports the covariance matrix for MOM (WIN-LOS) and VMG (VAL-GRO) for the three sets of months. The right-most covariance matrix is unconditional. In Panel B, we estimate the following time series regressions: $R_{p,t}^e = \alpha + \beta^0 R_{\mathrm{MKT,t}}^e + e_p$, where $R_{\mathrm{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with p = MOM (WIN-LOS) and VMG (VAL-GRO), and report the covariance of residual terms. In Panel C, we use the the Henriksson and Merton (1981) specification in each regression: $R_{p,t}^e = \alpha + \beta^0 R_{\mathrm{MKT,t}}^e + \beta^+ \max\left(R_{\mathrm{MKT,t}}^e, 0\right) + e_p$, where $R_{\mathrm{MKT,t}}^e$ is the market excess return and $R_{p,t}$ is the portfolio returns of our interest with p = MOM (WIN-LOS) and VMG (VAL-GRO), and report the covariance of residual terms. All numbers are reported in percentage squared per month.

PANEL A: COVARIANCE DECOMPOSITION OF RETURNS

	High				Med				Low				All		
	VMG	VAL	-GRO		VMG	VAL	-GRO		VMG	VAL	-GRO		VMG	VAL	-GRO
MOM	-64.05			MOM	-1.63			MOM	-1.12			MOM	-22.36		
WIN		66.45	-48.52	WIN		35.87	-28.40	WIN		19.46	-17.16	WIN		40.51	-31.37
-LOS		-174.14	92.16	-LOS		-34.56	25.46	-LOS		-19.47	16.05	-LOS		-75.92	44.41

PANEL B: COVARIANCE DECOMPOSITION OF CAPM RESIDUALS

	High				Med				Low				All		
	e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$
e_{MOM}	-35.58			e_{MOM}	-2.01			e_{MOM}	-1.21			e_{MOM}	-15.74		
e_{WIN}		-7.90	-1.61	e_{WIN}		1.88	-0.62	e_{WIN}		0.59	-0.25	e_{WIN}		-2.84	-0.96
$-e_{\mathrm{LOS}}$		-22.67	-3.39	$-e_{\mathrm{LOS}}$		-2.67	-0.60	$-e_{\mathrm{LOS}}$		-1.43	-0.12	$-e_{\mathrm{LOS}}$		-10.39	-1.54

PANEL C: COVARIANCE DECOMPOSITION OF HM RESIDUALS

	High				Med				Low				All		
	e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$		e_{VMG}	e_{VAL}	$-e_{\mathrm{GRO}}$
e_{MOM}	-28.83			e_{MOM}	-2.32			e_{MOM}	-1.14			e_{MOM}	-14.14		
e_{WIN}		-5.59	-1.37	e_{WIN}		1.67	-0.68	e_{WIN}		0.66	-0.28	e_{WIN}		-2.22	-0.94
$-e_{\mathrm{LOS}}$		-18.89	-2.99	$-e_{\mathrm{LOS}}$		-2.70	-0.61	$-e_{\mathrm{LOS}}$		-1.38	-0.14	$-e_{\mathrm{LOS}}$		-9.46	-1.51

Table 12: REGRESSION OF CAPM AND HM FITTED RESIDUALS ON LIQUIDITY SHOCKS

This table reports results of regressing the residuals from the CAPM and Henriksson-Merton (HM) specifications on proxies for economy-wide liquidity shocks. We consider (i) the funding liquidity shock measured by the innovations in AR(2) model of TED spread, the 3-month interbank LIBOR minus the 3- month T-bill rate, as well as (ii) the innovations in the aggregate market liquidity estimated by Pastor and Stambaugh (2003). We take the negative of the TED spread so that it repre- sents the level of liquidity. Panel A (B) reports regression results using CAPM (HM) residuals as regressands.

PANEL A: REGRESSION OF CAPM RESIDUALS ON LIQUIDITY RISK								
	High		Med		Low		All	
Liquidity Risk	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}
TED Spread	3.73	-3.86	-1.22	1.16	-1.01	-0.28	1.75	-1.97
	(1.40)	(-2.17)	(-0.45)	(0.75)	(-0.64)	(-0.21)	(1.09)	(-1.80)
Pastor-Stambaugh	9.40	-16.06	-10.63	5.50	-1.96	5.56	6.83	-8.47
	(0.97)	(-2.21)	(-1.29)	(0.90)	(-0.32)	(1.10)	(1.04)	(-1.99)

PANEL B: REGRESSION OF HM RESIDUALS ON LIQUIDITY RISK								
	High		Med		Low		All	
Liquidity Risk	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}	e_{MOM}	e_{VMG}
TED Spread	2.46	-2.78	-1.11	1.50	-1.54	-0.13	1.08	-1.60
	(0.91)	(-1.55)	(-0.40)	(0.99)	(-1.02)	(-0.09)	(0.70)	(-1.53)
Pastor-Stambaugh	2.95	-10.56	-11.13	3.92	-5.07	6.46	1.21	-5.33
	(0.31)	(-1.50)	(-1.37)	(0.65)	(-0.96)	(1.38)	(0.20)	(-1.31)

Table 13: Performance of Combined Value-Momentum Strategies

This table presents summary statistics for the monthly returns for five portfolios. MOM is the zero-investment momentum portfolio that is long the top-decile past winners and short the bottom-decile past losers. VMG is the zero-investment portfolio that is long the top-decile B/M stocks in short the bottom-decile B/M stocks. "50/50-static" is a portfolio with weights of 0.5 on the MOM and VMG portfolios. The "50/50-dyn" portfolio has weights of 0.5 on both the MOM and VMG portfolios when the turbulent state probability is in High tertile, and otherwise holds only the momentum portfolio. Finally, the " conditional MVE" puts weights of $w = (\gamma_s \boldsymbol{\Sigma}_s)^{-1} \boldsymbol{\mu}_s$ on VMG and MOM in each of two groups (High tertile and the rest) — that is, the portfolio in each group is the conditional *ex-post* mean variance efficient portfolio.

portfolio	mean	std	SR	skew.	kurt.	No. Rets < -10%	-15%	-20%
MOM	1.18	7.94	0.52	-2.43	21.22	56	32	13
VMG	0.52	6.46	0.28	2.62	24.96	27	9	3
50/50-static	0.85	3.87	0.76	-0.50	6.04	13	3	0
50/50-dyn	1.01	3.87	0.91	-0.33	4.45	8	2	0
Conditional MVE	1.18	3.87	1.05	0.35	5.33	8	1	0
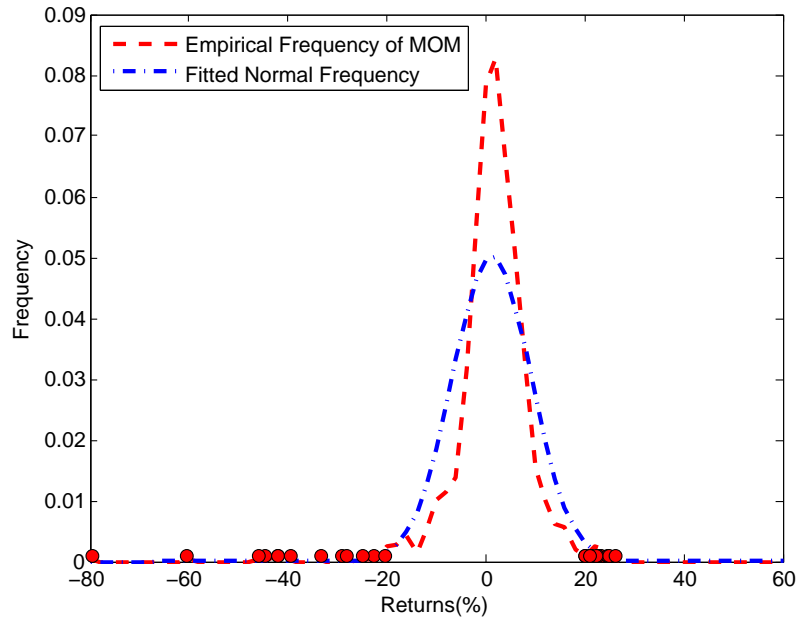
A Momentum Strategy Returns – Smoothed Density Function
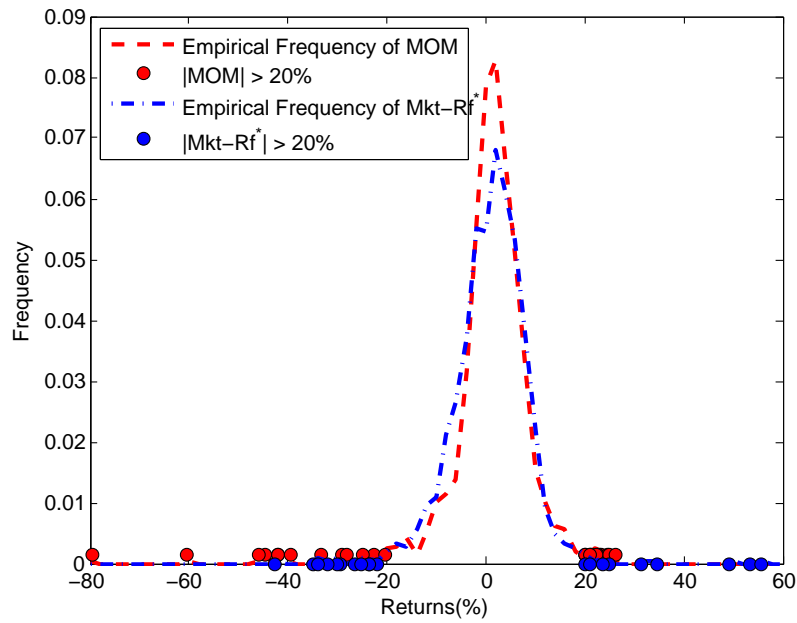
B Smoothed Density Functions–MOM and scaled excess market returns

Figure 1: EMPIRICAL FREQUENCY OF MOMENTUM STRATEGY RETURNS (MOM)

Panel A plots the smoothed empirical density of the MOM and the normal density with the same mean and standard deviation. To highlight the left skew of momentum strategy returns, we represent 25 MOM returns (13 in left tails and 12 in right tails) that exceed 20% in absolute value. Panel B plots the empirical density of MOM along with the empirical density of scaled market excess returns, Mkt-Rf*, with standard deviation equal to that of momentum strategy returns. The sample period is 1927:01-2013:12.
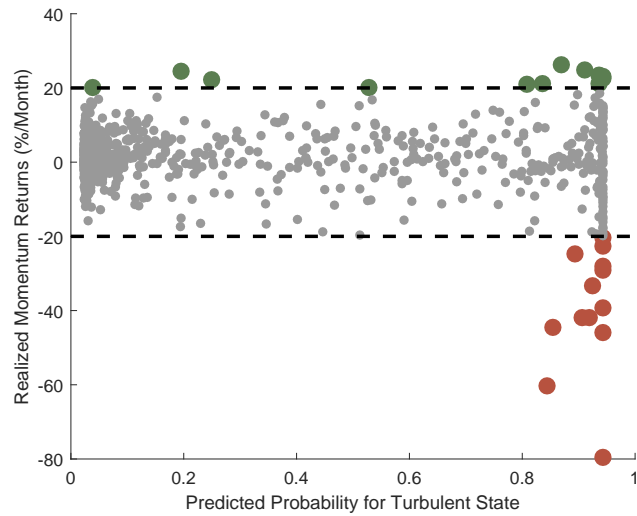
A In Sample Prediction
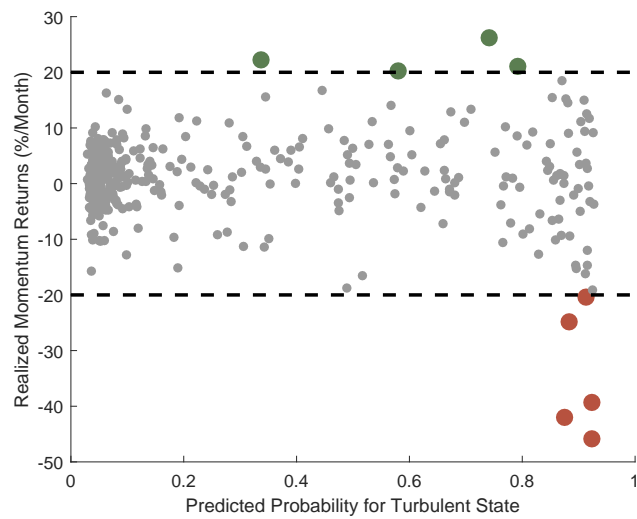
B Out of Sample Prediction

Figure 2: MOMENTUM RETURNS AND PROBABILITY OF THE HIDDEN STATE BEING TURBULENT

The figure presents a scatter plot of momentum strategy return on the vertical axis and $\Pr\left(S_t = \text{Turbulent}|\mathcal{F}_{t-1}\right)$, the probability that the hidden state is turbulent, on the horizontal axis. Momentum strategy returns below -20% are highlighted in red, and returns of exceeding 20% are in green. Figure (a) is based on in-sample estimates using all 1044 months (1927:01-2013:12). For each month t of the last 400 months in 1980:09-2013:12, we skip first 10 years over 1927:01-1936:12 and estimate our HMM using data from 1937:01 till month $t-1$ to compute $\Pr\left(S_t = \text{Turbulent}|\mathcal{F}_{t-1}\right)$. Figure (b) reports out-of-sample results.
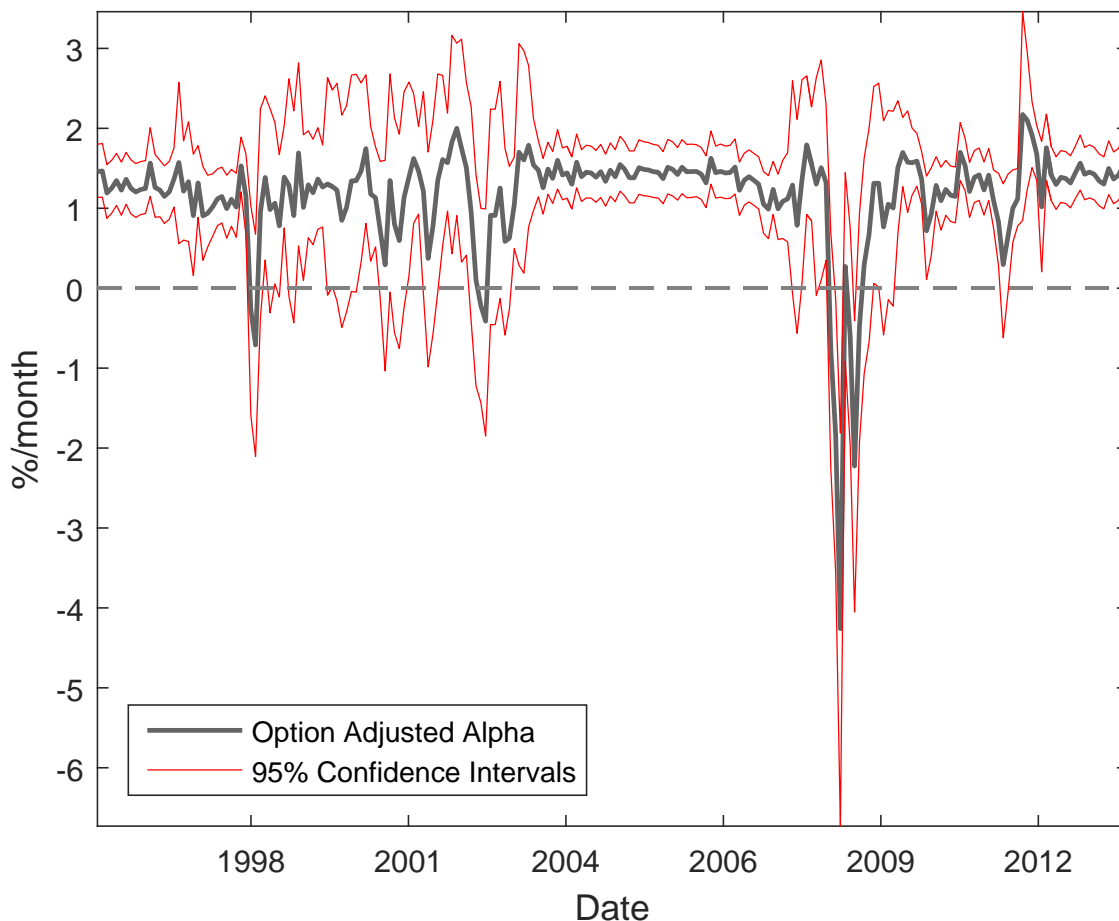
Figure 3: TIME SERIES OF OPTION ADJUSTED ALPHA

Option adjusted alpha, α^*, is computed by (16). The sample period is 1996:01 to 2013:12 where we can find the market price of call option on S&P 500 from OptionMetrics. 95% confidence intervals are computed as follows. First, we simulate 10,000 sets of parameters from the asymptotic distributions of parameters obtained from ML estimator. Then, for each set of parameters, we compute the monthly time series of α^*. Lastly, in each month, we find 95% confidence intervals of α^* by choosing top and bottom 2.5% quantiles from the simulated 10,000 observations of α^*.

A Hidden Markov Model of Leverage Dynamics, Tail Risk, and Value-Momentum Correlation

Online Appendices

Appendix A Inconsistency of QML

In many settings, it useful to assume that residuals are drawn from normal distributions in estimating a statistical model. When the true distribution of the residual is not normal, these estimates are Quasi-Maximum Likelihood (QML). Wooldridge (1986) provides sufficient conditions for the consistency and asymptotic normality of QML estimators. These conditions are not satisfied in our case. Below, we provide an example where the HMM return generating process innovations are drawn from a non-normal distribution and the resulting QML estimator—obtained by maximizing the misspecified normal likelihood—gives an asymptotically biased (inconsistent) estimate of the true parameter value.

Suppose R_t follows the process given below:

$$R_t = \sigma\left(S_t\right)\varepsilon_t, \tag{A.1}$$

where $\sigma\left(S_t\right)$ is either σ_{H} or σ_{L}, depending on the realization of hidden state of S_t which is either H or L. The transition probability matrix that determines the evolution of the hidden state S_t is given by

$$\Pi = \left[\begin{array}{cc} \Pr(S_t = \mathrm{H}|S_{t-1} = \mathrm{H}) & \Pr(S_t = \mathrm{L}|S_{t-1} = \mathrm{H}) \\ \Pr(S_t = \mathrm{H}|S_{t-1} = \mathrm{L}) & \Pr(S_t = \mathrm{L}|S_{t-1} = \mathrm{L}) \end{array}\right] = \left[\begin{array}{cc} p & 1-p \\ 1-p & p \end{array}\right]. \tag{A.2}$$

An econometrician observes the time series of $\{R_t\}_{t=1}^{T}$ but not the underlying state. The parameters p and σ_{L} are known. The econometrician estimates the unknown parameter σ_{H} by QML, that is by assuming that ε_t is drawn from the standard normal distribution, whereas ε_t is either 1 or -1 with equal probability. In what follows, we show that when

$$\sigma_{\mathrm{H}} = 1.5, \sigma_{\mathrm{L}} = 1, \text{ and } p = 0.52, \tag{A.3}$$

the QML estimator of σ_{H} is inconsistent.

The misspecified normal log likelihood of $\{R_t\}_{t=1}^{T}$ is given by

$$\frac{1}{T}\sum_{t=1}^{T}\log\left(\mathcal{L}\left(R_t\right)\right), \tag{A.4}$$

where

$$\mathcal{L}\left(R_t\right) = \lambda_{t-1}\phi\left(R_t|\sigma_{\mathrm{H}}\right) + \left(1 - \lambda_{t-1}\right)\phi(R_t|\sigma_{\mathrm{L}}), \tag{A.5}$$

$\phi\left(x|\sigma\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)$ is the density function of $\mathcal{N}\left(0,\sigma^2\right)$, and λ_{t-1} is the probability for $S_t = H$ given the information set $\mathcal{F}_{t-1} = \{R_1, R_2, \cdots, R_{t-1}\}$ when the econometrician uses the (incorrect) normal density for inference. When the true likelihood is used, let λ_{t-1}^* denote the probability of $S_t = H$ given $\mathcal{F}_{t-1}$. Since S_t is hidden, both λ_{t-1} and λ_{t-1}^* are weighted averages of

p and $1 - p$ and the following should be satisfied:

$$1 - p \leq \lambda_{t-1}, \lambda_{t-1}^* \leq p \qquad (A.6)$$

for every $\mathcal{F}_{t-1}$.

The QML estimate $\widehat{\sigma}_{\mathrm{H}}$ is obtained by maximizing (A.4), giving rise to the first order condition:

$$\frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \bigg|_{\sigma_{\mathrm{H}} = \widehat{\sigma}_{\mathrm{H}}} = 0. \qquad (A.7)$$

If $\widehat{\sigma}_{\mathrm{H}}$ converges to σ_{H}^0, the LHS of (A.7) converges to the true expectation as

$$\frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \bigg|_{\sigma_{\mathrm{H}} = \widehat{\sigma}_{\mathrm{H}}} \xrightarrow{p} \mathbb{E} \left[\frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \right]_{\sigma_{\mathrm{H}} = \sigma_{\mathrm{H}}^0} \qquad (A.8)$$

under mild regularity conditions. Noting that the RHS of (A.7) is always zero, it follows that

$$\mathbb{E} \left[\frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \right]_{\sigma_{\mathrm{H}} = \sigma_{\mathrm{H}}^0} = \mathbb{E} \left[\mathbb{E} \left[\frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \big| \mathcal{F}_{t-1} \right] \right]_{\sigma_{\mathrm{H}} = \sigma_{\mathrm{H}}^0} = 0. \qquad (A.9)$$

We show the inconsistency of $\widehat{\sigma}_{\mathrm{H}}$ by verifying that (A.9) cannot hold. When $\sigma_{\mathrm{H}} = \sigma_{\mathrm{H}}^0$, there exists $\delta > 0$ such that $\mathbb{E} \left[\frac{\partial \log (\mathcal{L} (R_t))}{\partial \sigma_{\mathrm{H}}} | \mathcal{F}_{t-1} \right] < -\delta$ for every $\mathcal{F}_{t-1}$, implying that $\mathbb{E} \left[\frac{\partial \log (\mathcal{L} (R_t | \sigma_{\mathrm{H}}))}{\partial \sigma_{\mathrm{H}}} \right] < -\delta$.

Hereafter, we will evaluate the conditional expectation at $\sigma_{\mathrm{H}} = \sigma_{\mathrm{H}}^0$. From (A.5), note that $\mathbb{E} \left[\frac{\partial \log (\mathcal{L} (R_t))}{\partial \sigma_{\mathrm{H}}} | \mathcal{F}_{t-1} \right]$ is decomposed as follows:

$$\begin{aligned}
\mathbb{E} \left[\frac{\partial \log \left(\mathcal{L} \left(R_t \right) \right)}{\partial \sigma_{\mathrm{H}}} \big| \mathcal{F}_{t-1} \right] &= \mathbb{E} \left[\frac{\lambda_{t-1}}{\mathcal{L} \left(R_t \right)} \frac{\partial \phi \left(R_t | \sigma_{\mathrm{H}} \right)}{\partial \sigma_{\mathrm{H}}} \big| \mathcal{F}_{t-1} \right] \\
&+ \mathbb{E} \left[\frac{1}{\mathcal{L} \left(R_t \right)} \left(\phi \left(R_t | \sigma_{\mathrm{H}} \right) - \phi \left(R_t | \sigma_{\mathrm{L}} \right) \right) | \mathcal{F}_{t-1} \right] \frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}}. \qquad (A.10)
\end{aligned}$$

To determine the sign of each component in RHS of (A.10), we need the conditional distribution of R_t. Since λ_{t-1}^* is the true probability of $S_t = H$ given $\mathcal{F}_{t-1}$ and ε_t in (A.1) is drawn from a binomial distribution of 1 or -1 with equal probability, the probability mass of R_t over $(-\sigma_{\mathrm{H}}, -\sigma_{\mathrm{L}}, \sigma_{\mathrm{L}}, \sigma_{\mathrm{H}})$ equals $\left(\frac{\lambda_{t-1}^*}{2}, \frac{1 - \lambda_{t-1}^*}{2}, \frac{1 - \lambda_{t-1}^*}{2}, \frac{\lambda_{t-1}^*}{2} \right)$.

First, we determine the sign of $\mathbb{E} \left[\frac{\lambda_{t-1}}{\mathcal{L}(R_t)} \frac{\partial \phi(R_t | \sigma_{\mathrm{H}})}{\partial \sigma_{\mathrm{H}}} | \mathcal{F}_{t-1} \right]$. From the properties of the normal

density, it follows that $\frac{\partial \phi(x|\sigma)}{\partial \sigma} = \phi(x|\sigma)\left(-\frac{1}{\sigma} + \frac{x^2}{\sigma^3}\right)$ and $\phi(-x|\sigma) = \phi(x|\sigma)$. Hence

$$
\begin{aligned}
\mathbb{E}\left[\frac{\lambda_{t-1}}{\mathcal{L}}\frac{\partial \phi\left(R_t|\sigma_{\mathrm{H}}\right)}{\partial \sigma_{\mathrm{H}}}\Big|\mathcal{F}_{t-1}\right] &= \frac{\lambda_{t-1}^*}{2}\sum_{R_t=-\sigma_{\mathrm{H}},\sigma_{\mathrm{H}}}\frac{\lambda_{t-1}}{\mathcal{L}\left(R_t\right)}\phi(R_t|\sigma_{\mathrm{H}})\left(-\frac{1}{\sigma_{\mathrm{H}}}+\frac{R_t^2}{\sigma_{\mathrm{H}}^3}\right) \\
&+ \frac{1-\lambda_{t-1}^*}{2}\sum_{R_t=-\sigma_{\mathrm{L}},\sigma_{\mathrm{L}}}\frac{\lambda_{t-1}}{\mathcal{L}\left(R_t\right)}\phi(R_t|\sigma_{\mathrm{H}})\left(-\frac{1}{\sigma_{\mathrm{H}}}+\frac{R_t^2}{\sigma_{\mathrm{H}}^3}\right) \\
&= \frac{\left(1-\lambda_{t-1}^*\right)\lambda_{t-1}}{\mathcal{L}(\sigma_{\mathrm{L}})}\phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}})\left(-\frac{1}{\sigma_{\mathrm{H}}}+\frac{\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{H}}^3}\right) \\
&< -(1-p)^2\frac{\phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}})}{\phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}})}\left(\frac{\sigma_{\mathrm{H}}^2-\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{H}}^3}\right),
\end{aligned}
\tag{A.11}
$$

where the last inequality is from (A.6) and $\mathcal{L}(\sigma_{\mathrm{L}}) < \phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}})$.

Next, from the property, $\phi(-x|\sigma) = \phi(x|\sigma)$, and the fact that $\phi\left(x|\sigma\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{x^2}{2\sigma^2}\right)$, the sign of $\mathbb{E}\left[\frac{1}{\mathcal{L}}\left(\phi\left(R_t|\sigma_{\mathrm{H}}\right) - \phi\left(R_t|\sigma_{\mathrm{L}}\right)\right)|\mathcal{F}_{t-1}\right]$ is determined as follows:

$$
\begin{aligned}
&\mathbb{E}\left[\frac{1}{\mathcal{L}}\left(\phi\left(R_t|\sigma_{\mathrm{H}}\right) - \phi\left(R_t|\sigma_{\mathrm{L}}\right)\right)|\mathcal{F}_{t-1}\right] \\
&= \frac{\lambda_{t-1}^*}{2}\sum_{R_t=-\sigma_{\mathrm{H}},\sigma_{\mathrm{H}}}\left(\frac{\phi\left(R_t|\sigma_{\mathrm{H}}\right)-\phi\left(R_t|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(R_t\right)}\right) + \frac{1-\lambda_{t-1}^*}{2}\sum_{R_t=-\sigma_{\mathrm{L}},\sigma_{\mathrm{L}}}\left(\frac{\phi\left(R_t|\sigma_{\mathrm{H}}\right)-\phi\left(R_t|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(R_t\right)}\right) \\
&= \lambda_{t-1}^*\left(\frac{\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}\right) + \left(1-\lambda_{t-1}^*\right)\left(\frac{\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(\sigma_{\mathrm{L}}\right)}\right) \\
&= \lambda_{t-1}^*\left(\frac{\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}\right) + \left(1-\lambda_{t-1}^*\right)\frac{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}{\mathcal{L}\left(\sigma_{\mathrm{L}}\right)}\left(\frac{\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}}\right)}{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}\right) \\
&> \frac{1}{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}\left(\lambda_{t-1}^*\left(\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{L}}\right)\right) + \left(1-\lambda_{t-1}^*\right)\left(\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}}\right)\right)\right) \\
&> \frac{1}{\mathcal{L}\left(\sigma_{\mathrm{H}}\right)}\left((1-p)\left(\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{H}}|\sigma_{\mathrm{L}}\right)\right) + p\left(\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}}\right)-\phi\left(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}}\right)\right)\right) > 0,
\end{aligned}
\tag{A.12}
$$

where the last three inequalities can be verified by (A.6) and the given parameter values of (A.3).

Finally, we show that $\frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}} \leq 0$ by induction. We assume that λ_0 is determined as the steady state distribution determined by (A.2). Since λ_0 does not depend on σ_{H}, the following holds:

$$
\frac{\partial \lambda_0}{\partial \sigma_{\mathrm{H}}} = 0.
\tag{A.13}
$$

Next, we show that $\frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}} \leq 0$ implies $\frac{\partial \lambda_t}{\partial \sigma_{\mathrm{H}}} \leq 0$. Note that the process of $\{\lambda_t\}_{t=0}^T$ is constructed by the following recursion:

$$
\widetilde{\lambda}_t = \frac{\lambda_{t-1}\phi\left(R_t|\sigma_{\mathrm{H}}\right)}{\lambda_{t-1}\phi\left(R_t|\sigma_{\mathrm{H}}\right) + (1-\lambda_{t-1})\phi\left(R_t|\sigma_{\mathrm{L}}\right)},
\tag{A.14}
$$

and

$$
\lambda_t = p\widetilde{\lambda}_t + (1-p)\left(1-\widetilde{\lambda}_t\right).
\tag{A.15}
$$

Equation (A.14) describes how the econometrician updates the probability on the hidden state of S_t using the misspecified normal likelihood after observing R_t. Equation (A.15) shows how the econometrician predicts the hidden state of S_{t+1} with the given information set $\mathcal{F}_t$ through the transition matrix given in (A.2). Combining (A.14) and (A.15), we get

$$\frac{\lambda_t + p - 1}{2p - 1} = \frac{\lambda_{t-1} \phi\left(R_t | \sigma_{\mathrm{H}}\right)}{\lambda_{t-1} \phi\left(R_t | \sigma_{\mathrm{H}}\right) + \left(1 - \lambda_{t-1}\right) \phi\left(R_t | \sigma_{\mathrm{L}}\right)}. \tag{A.16}$$

Taking the derivative of (A.16) with respect to σ_{H}, we obtain the following:

$$\frac{1}{2p - 1} \frac{\partial \lambda_t}{\partial \sigma_{\mathrm{H}}} = \frac{\partial \frac{\lambda_{t-1} \phi(R_t | \sigma_{\mathrm{H}})}{\lambda_{t-1} \phi(R_t | \sigma_{\mathrm{H}}) + (1 - \lambda_{t-1}) \phi(R_t | \sigma_{\mathrm{L}})}}{\partial \lambda_{t-1}} \frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}} + \frac{\partial \frac{\lambda \phi(R_t | \sigma_{\mathrm{H}})}{\lambda \phi(R_t | \sigma_{\mathrm{H}}) + (1 - \lambda) \phi(R_t | \sigma_{\mathrm{L}})}}{\partial \phi\left(R_t | \sigma_{\mathrm{H}}\right)} \frac{\partial \phi\left(R_t | \sigma_{\mathrm{H}}\right)}{\partial \sigma_{\mathrm{H}}}. \tag{A.17}$$

To determine the sign of each component in RHS of (A.17), we use the following properties:

$$\frac{\partial \frac{\lambda m}{\lambda m + (1 - \lambda) n}}{\partial \lambda} = \frac{mn}{\left(\lambda m + (1 - \lambda) n\right)^2} > 0 \tag{A.18}$$

$$\frac{\partial \frac{\lambda m}{\lambda m + (1 - \lambda) n}}{\partial m} = \frac{\lambda(1 - \lambda) n}{\left(\lambda m + (1 - \lambda) n\right)^2} > 0 \tag{A.19}$$

for $m, n > 0$ and $\lambda \in (0, 1)$. Further, using the properties of $\frac{\partial \phi(x|\sigma)}{\partial \sigma} = \phi(x|\sigma)\left(-\frac{1}{\sigma} + \frac{x^2}{\sigma^3}\right)$ and $\phi\left(x|\sigma\right) = \phi\left(-x|\sigma\right)$, we have that

$$\frac{\partial \phi\left(\sigma_{\mathrm{H}} | \sigma_{\mathrm{H}}\right)}{\partial \sigma_{\mathrm{H}}} = \phi(\sigma_{\mathrm{H}} | \sigma_{\mathrm{H}})\left(-\frac{1}{\sigma_{\mathrm{H}}} + \frac{\sigma_{\mathrm{H}}^2}{\sigma_{\mathrm{H}}^3}\right) = 0$$

$$\frac{\partial \phi\left(\sigma_{\mathrm{L}} | \sigma_{\mathrm{H}}\right)}{\partial \sigma_{\mathrm{H}}} = \phi(\sigma_{\mathrm{L}} | \sigma_{\mathrm{H}})\left(-\frac{1}{\sigma_{\mathrm{H}}} + \frac{\sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{H}}^3}\right) < 0,$$

implying

$$\frac{\partial \phi\left(R_t | \sigma_{\mathrm{H}}\right)}{\partial \sigma_{\mathrm{H}}} \leq 0 \tag{A.20}$$

for every possible realization of R_t from $\{-\sigma_{\mathrm{H}}, -\sigma_{\mathrm{L}}, \sigma_{\mathrm{L}}, \sigma_{\mathrm{H}}\}$. With the assumption that $\frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}} \leq 0$, inequalities of (A.18), (A.19), and (A.20) ensure that RHS of (A.17) is non-positive. Hence, with $p > 1/2$ as assumed in (A.3), it follows that $\frac{\partial \lambda_t}{\partial \sigma_{\mathrm{H}}} \leq 0$. Combining (A.13) with this finding, we conclude that

$$\frac{\partial \lambda_{t-1}}{\partial \sigma_{\mathrm{H}}} \leq 0, \tag{A.21}$$

for every possible information set of $\mathcal{F}_{t-1}$.

Recall that we want to show that (A.10) is strictly negative. Finally, combining (A.11), (A.12), and (A.20), we conclude that

$$\mathbb{E}\left[\frac{\partial \log\left(\mathcal{L}\left(R_t | \sigma_{\mathrm{H}}\right)\right)}{\partial \sigma_{\mathrm{H}}} | \mathcal{F}_{t-1}\right] < -\delta, \tag{A.22}$$

iv

where

$$\delta = (1-p)^2 \frac{\phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{H}})}{\phi(\sigma_{\mathrm{L}}|\sigma_{\mathrm{L}})} \left(\frac{\sigma_{\mathrm{H}}^2 - \sigma_{\mathrm{L}}^2}{\sigma_{\mathrm{H}}^3} \right) > 0, \tag{A.23}$$

completing the proof that QML estimate of $\widehat{\sigma}_{\mathrm{H}}$ in (A.7) will not converge to the true parameter value.

Appendix B   Additional Tables

Table 14: Option-like Feature of Momentum Returns and Market Conditions

We partition the months in our sample into three groups: 'High' group is made up of months when variable describing the market conditions (past market returns, realized volatility of the market, or leverage of loser portfolio stocks) was in the top 20th percentile and the 'Low' group corresponds to months when the market condition variable was in the bottom 20th percentile. The rest of the months are classified as 'Medium'. The sample period is 1927:07-2013:12. In Panel A, we group the months based on the realized volatility of daily market returns over the previous 12 months. In Panel B, we use the breakpoints of the loser portfolio for grouping. We then pool the months within each group and analyze the behavior of momentum strategy returns. Specifically, we estimate equation (1) with ordinary least squares using momentum strategy returns (R_{MOM}) and the returns of winner and loser portfolio in excess of risk free return (R_{WIN}^e and R_{LOS}^e) as LHS variables and report results in Panel A-1-I and B-1-I. For comparison, we report the estimates for the CAPM, without the exposure to the call option on the market in (1), in Panel A-1-II and B-1-II. Then, we count the numbers of large momentum losses worse than negative 20% within the groups and report those in Panel A-2 and B-2. Finally, we compare the skewness of $R_{p,t}^e$ with that of estimated ε of (1) in Panel A-3 and B-3. α is reported in percentage per month. The t-statistics are computed using the heteroscedasticity-consistent covariance estimator by White (1980).

Table 14 – continued from previous page

	PANEL A: PAST 12 MONTHS REALIZED VOLATILITY OF MARKET RETURNS								
	HIGH			MEDIUM			LOW		
LHS	R_{MOM}	R^e_{WIN}	R^e_{LOS}	R_{MOM}	R^e_{WIN}	R^e_{LOS}	R_{MOM}	R^e_{WIN}	R^e_{LOS}
	A-1: OPTION-LIKE FEATURES								
	A-1-I: HENRIKSSON-MERTON ESTIMATES								
α	2.90	1.07	-1.83	1.93	0.77	-1.16	2.40	1.38	-1.02
$t(\alpha)$	(2.96)	(2.71)	(-2.58)	(5.73)	(4.05)	(-5.01)	(4.55)	(4.21)	(-2.98)
β^0	-0.59	0.94	1.52	0.16	1.35	1.19	0.54	1.55	1.02
$t(\beta^0)$	(-4.83)	(13.78)	(17.78)	(1.72)	(25.36)	(18.24)	(3.00)	(14.91)	(8.23)
β^+	-0.91	-0.27	0.63	-0.25	-0.19	0.06	-0.63	-0.46	0.17
$t(\beta^+)$	(-3.23)	(-2.14)	(3.39)	(-1.38)	(-1.93)	(0.51)	(-1.92)	(-2.39)	(0.79)
$Adj.R^2(\%)$	0.49	0.74	0.83	0.00	0.78	0.68	0.03	0.73	0.57
	A-1-II: CAPM ESTIMATES								
α	0.12	0.23	0.11	1.48	0.43	-1.04	1.58	0.78	-0.80
$t(\alpha)$	(0.18)	(0.82)	(0.20)	(6.69)	(3.58)	(-6.96)	(4.87)	(4.16)	(-3.56)
β	-1.10	0.78	1.88	0.05	1.27	1.22	0.19	1.30	1.11
$t(\beta)$	(-8.43)	(14.68)	(21.61)	(0.78)	(41.31)	(29.99)	(1.83)	(23.47)	(15.56)
$Adj.R^2$	0.45	0.73	0.82	0.00	0.78	0.68	0.01	0.72	0.57
	A-2: NUMBER OF MOMENTUM LOSSES WORSE THAN -20%								
	13			0			0		
	A-3: CONDITIONAL SKEWNESS								
LHS	-1.88	-0.21	1.42	-0.17	-0.65	-0.23	0.00	-0.13	0.16
ε	-0.62	-0.86	0.70	-0.11	0.33	0.41	-0.01	0.59	0.48

Table 14 – continued from previous page

Panel B: Breakpoints of Loser Portfolio									
		Low			Medium			High	
LHS	R_{MOM}	R^e_{WIN}	R^e_{LOS}	R_{MOM}	R^e_{WIN}	R^e_{LOS}	R_{MOM}	R^e_{WIN}	R^e_{LOS}
B-1: Option-like features									
B-1-i: Henriksson-Merton Estimates									
α	2.67	0.96	-1.71	2.79	1.21	-1.58	0.81	0.32	-0.50
$t(\alpha)$	(2.67)	(2.35)	(-2.38)	(5.82)	(6.23)	(-4.50)	(1.40)	(0.84)	(-1.51)
β^0	-0.65	0.91	1.56	0.22	1.39	1.17	0.52	1.48	0.96
$t(\beta^0)$	(-5.46)	(14.09)	(17.98)	(1.83)	(25.31)	(13.89)	(2.96)	(10.89)	(13.27)
β^+	-0.92	-0.29	0.63	-0.61	-0.35	0.26	-0.14	-0.09	0.05
$t(\beta^+)$	(-3.31)	(-2.37)	(3.34)	(-2.07)	(-3.18)	(1.23)	(-0.42)	(-0.44)	(0.27)
$Adj.R^2$	0.50	0.70	0.83	0.03	0.80	0.67	0.16	0.81	0.75
B-1-ii: CAPM Estimates									
α	-0.07	0.10	0.16	1.76	0.62	-1.14	0.57	0.17	-0.40
$t(\alpha)$	(-0.09)	(0.33)	(0.31)	(8.50)	(5.77)	(-7.58)	(1.86)	(0.76)	(-2.50)
β	-1.15	0.75	1.91	-0.08	1.22	1.30	0.45	1.43	0.98
$t(\beta)$	(-9.05)	(14.61)	(22.14)	(-0.95)	(34.29)	(20.93)	(4.67)	(25.29)	(16.17)
$Adj.R^2$	0.47	0.69	0.82	0.00	0.79	0.67	0.17	0.81	0.75
B-2: Number of Momentum Losses worse than -20%									
		12			1			0	
B-3: Conditional Skewness									
LHS	-1.70	-0.02	1.44	-1.21	-0.73	0.50	0.04	-0.51	0.07
ε	-0.39	0.06	0.75	-0.72	-0.05	0.69	-0.09	0.31	0.70

Table 15: Conditional covariance of momentum and value factor returns

This table presents the conditional covariance of MOM with three value factors: the HML factor by FF (Fama and French, 1993) and two value factors by AMP (Asness et al., 2013) – i) "Value Everywhere" which utilizes all assets across many markets and countries and ii) "Value US Equity" which uses only assets in US equity market. We group the months in our sample into three equal-sized tertiles (High, Med, Low) based on the predicted probability of the hidden state being turbulent, $\Pr\left(S_t = Turbulent | \mathcal{F}_{t1}\right)$. All numbers are reported in percentage squared per month.

Value Factor	Sample Period	High	Med	Low	All
HML(FF)	1927:01-2013:12 (1044 months)	-29.97	-1.20	-0.46	-10.51
Value Everywhere (AMP)	1972:01-2013:12 (504 months)	-14.16	-2.60	-1.32	-6.14
Value US equity (AMP)	1972:02-2013:12 (503 months)	-31.96	-7.31	-3.07	-14.20