Tail Risk in Momentum Strategy Returns

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:

$$1-p \leq \lambda_{t-1}, \lambda^*_{t-1} \leq p \tag{A.6}$$

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:

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

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

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

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

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

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:

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

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

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

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 $\phi(x|\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$, 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(\left(1-p\right)\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, \qquad \text{(A.12)}
\end{aligned}
$$

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:

$$
\frac{\partial\lambda_0}{\partial\sigma_{\mathrm{H}}} = 0. \qquad \text{(A.13)}
$$

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:

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

and

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

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

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

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

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

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

$$\frac{\partial \frac{\lambda m}{\lambda m+(1-\lambda)n}}{\partial \lambda} = \frac{mn}{(\lambda m+(1-\lambda)n)^2} > 0 \tag{A.18}$$
$$\frac{\partial \frac{\lambda m}{\lambda m+(1-\lambda)n}}{\partial m} = \frac{\lambda(1-\lambda)n}{(\lambda m+(1-\lambda)n)^2} > 0 \tag{A.19}$$

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 $\phi(x|\sigma) = \phi(-x|\sigma)$, we have that

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

implying

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

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

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

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'. For Panel A, the sample period is 1929:07-2013:12. For Panel B and C, the sample period is 1927:07-2013:12. In Panel A, we group the sample on the basis of cumulative market return during the 36 months preceding the month in which the momentum portfolios are formed. In Panel B, we group the months based on the realized volatility of daily market returns over the previous 12 months. In Panel C, 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^e_{WIN} and R^e_{LOS}) as LHS variables and report results in Panel A-1-i, B-1-i, and C-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, B-1-ii, and C-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, B-2, and C-2. Finally, we compare the skewness of $R^e_{p,t}$ with that of estimated ε of (1) in Panel A-3, B-3, and C-3. α is reported in percentage per month. The t-statistics are computed using the heteroscedasticity-consistent covariance estimator by White (1980).

Continued on next page

Table 14 – continued from previous page

Panel B: 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}
B-1: Option-like features									
B-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
B-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
B-2: Number of Momentum Losses worse than -20%									
	13			0			0		
B-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

Continued on next page

Panel C: 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}
C-1: Option-like features									
C-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
C-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
C-2: Number of Momentum Losses worse than -20%									
	12			1			0		
C-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, $\Pr(S_t = Turbulent|\mathcal{F}_{t1})$. 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